

10671921

4/8/04

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
15 November 2001 (15.11.2001)

PCT

(10) International Publication Number
WO 01/85942 A2

(51) International Patent Classification⁷: **C12N 15/12**,
15/10, C12Q 1/68, C07K 14/47, 14/705, 16/18, 16/28,
A01K 67/027, A61K 38/17, 39/395, G01N 33/53, 33/577

(21) International Application Number: PCT/US01/14355

(22) International Filing Date: 3 May 2001 (03.05.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/201,960 5 May 2000 (05.05.2000) US
60/202,729 8 May 2000 (08.05.2000) US
60/209,705 5 June 2000 (05.06.2000) US
60/210,149 7 June 2000 (07.06.2000) US
60/213,215 21 June 2000 (21.06.2000) US

(71) Applicant (for all designated States except US): **INCYTE GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **TANG, Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 94304 (US). **AU-YOUNG, Janice** [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US). **LU, Dyung, Aina, M.** [US/US]; 233 Coy Drive, San Jose, CA 95123 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **HILLMAN, Jennifer, L.** [US/US]; 230 Monroe Drive #17, Mountain View, CA 94040 (US). **AZIMZAI, Yalda** [US/US]; 5518 Boulder Canyon Drive, Castro Valley, CA 94552 (US).

LAL, Preeti [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). **YAO, Monique, G.** [US/US]; 111 Frederick Court, Mountain View, CA 94043 (US). **BANDMAN, Olga** [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). **BURFORD, Neil** [US/US]; 105 Wildwood Circle, Durham, CT 06422 (US). **BATRA, Sajeev** [US/US]; 555 El Camino Real #709, Sunnyvale, CA 94087 (US). **KEARNEY, Liam** [IE/US]; 50 Woodside Avenue, San Francisco, CA 94127 (US). **POLICKY, Jennifer, L.** [US/US]; 1511 Jarvis Court, San Jose, CA 95118 (US).

(74) Agents: **HAMLET-COX, Diana et al.**; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/85942 A2

Applicants: Paz Einat et al.
Serial No.: 10/671,921
Filed: September 24, 2003
Exhibit 4

(54) Title: CYTOSKELETON-ASSOCIATED PROTEINS

(57) Abstract: The invention provides human cytoskeleton-associated proteins (CYSKP) and polynucleotides which identify and encode CYSKP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of CYSKP.

CYTOSKELETON-ASSOCIATED PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of cytoskeleton-associated proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative, autoimmune/inflammatory, vesicle trafficking, neurological, cell motility, reproductive, and muscle disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of cytoskeleton-associated proteins.

BACKGROUND OF THE INVENTION

The cytoskeleton, a cytoplasmic system of protein fibers, mediates cell shape, structure, and movement. The cytoskeleton supports the cell membrane and forms tracks along which organelles and other elements move in the cytosol. The cytoskeleton is a dynamic structure that allows cells to adopt various shapes and to carry out directed movements. Additionally, molecules can be sequestered to a specific cellular location through interaction with cytoskeleton associated proteins. Major cytoskeletal fibers are the microfilaments, the microtubules, and the intermediate filaments. Motor proteins, including myosin, dynein, and kinesin, drive movement of, or along, the fibers. Accessory or associated proteins modify the structure or activity of the fibers while cytoskeletal membrane anchors connect the fibers to the cell membrane. Other proteins associated with the cytoskeleton have roles in processes such as secretion and intracellular signaling. (The cytoskeleton is reviewed in Lodish, H. et al. (1995) Molecular Cell Biology Scientific American Books, New York NY.)

Microtubules and Associated Proteins

Tubulins

Microtubules, cytoskeletal fibers with a diameter of 24 nm, have multiple roles in the cell. Bundles of microtubules form cilia and flagella, which are whip-like extensions of the cell membrane that are necessary for sweeping materials across an epithelium and for swimming of sperm, respectively. Marginal bands of microtubules in red blood cells and platelets are important for these cells' pliability. Organelles, membrane vesicles, and proteins are transported in the cell along tracks of microtubules. For example, microtubules run through nerve cell axons, allowing bi-directional transport of materials and membrane vesicles between the cell body and the nerve terminal. Failure to supply the nerve terminal with these vesicles blocks the transmission of neural signals. Microtubules, in the form of the spindle, are also critical to chromosomal movement during cell division. Both stable and short-lived populations of microtubules exist in the cell.

Microtubules are a polymer of GTP-binding tubulin protein subunits. Each subunit is a

heterodimer of α - and β - tubulin, multiple isoforms of which exist. Alpha-tubulin undergoes a number of post-translational modifications, including acetylation, polyglutamylation, truncation of two amino acids (forming $\Delta 2$ tubulin), and tyrosination. In some cases, these modifications can affect microtubule stability. The hydrolysis of GTP is linked to the addition of tubulin subunits at the end of a
5 microtubule. The subunits interact head to tail to form protofilaments; the protofilaments interact side to side to form a microtubule. A microtubule is polarized, one end ringed with α -tubulin and the other with β -tubulin, and the two ends differ in their rates of assembly. Each microtubule is generally composed of 13 protofilaments although 11 or 15 protofilament-microtubules are sometimes found. Cilia and flagella contain doublet microtubules. Microtubules grow from specialized structures known
10 as centrosomes or microtubule-organizing centers (MTOCs). MTOCs may contain one or two centrioles, which are pinwheel arrays of triplet microtubules. The basal body, the organizing center located at the base of a cilium or flagellum, contains one centriole. γ - tubulin present in the MTOC is important for nucleating the polymerization of α - and β - tubulin heterodimers but does not polymerize into microtubules. The protein pericentrin is found in the MTOC and has a role in microtubule
15 assembly.

Microtubule-Associated Proteins

Microtubule-associated proteins (MAPs) have roles in the assembly and stabilization of microtubules. One major family of MAPs, assembly MAPs, can be identified in neurons as well as non-neuronal cells. Assembly MAPs are responsible for cross-linking microtubules in the cytosol.
20 These MAPs are organized into two domains: a basic microtubule-binding domain and an acidic projection domain. The projection domain is the binding site for membranes, intermediate filaments, or other microtubules. Based on sequence analysis, assembly MAPs can be further grouped into two types: Type I and Type II.

Type I MAPs, which include MAP1A and MAP1B, are large, filamentous molecules that co-
25 purify with microtubules and are abundantly expressed in brain and testis. They contain several repeats of a positively-charged amino acid sequence motif that binds and neutralizes negatively charged tubulin, leading to stabilization of microtubules. MAP1A and MAP1B are each derived from a single precursor polypeptide that is subsequently proteolytically processed to generate one heavy chain and one light chain.

30 Another light chain, LC3, is a 16.4 kDa molecule that binds MAP1A, MAP1B, and microtubules. It is suggested that LC3 is synthesized from a source other than the MAP1A or MAP1B transcripts, and the expression of LC3 may be important in regulating the microtubule binding activity of MAP1A and MAP1B during cell proliferation (Mann, S. S. et al. (1994) J. Biol. Chem. 269:11492-11497).

Type II MAPs, which include MAP2a, MAP2b, MAP2c, MAP4, and Tau, are characterized by three to four copies of an 18-residue sequence in the microtubule-binding domain. MAP2a, MAP2b, and MAP2c are found only in dendrites, MAP4 is found in non-neuronal cells, and Tau is found in axons and dendrites of nerve cells. Alternative splicing of the Tau mRNA leads to the existence of multiple forms of Tau protein. Tau phosphorylation is altered in neurodegenerative disorders such as Alzheimer's disease, Pick's disease, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia and Parkinsonism linked to chromosome 17. The altered Tau phosphorylation leads to a collapse of the microtubule network and the formation of intraneuronal Tau aggregates (Spillantini, M.G. and Goedert, M. (1998) Trends Neurosci. 21:428-433).

Microtubule stability may also be regulated by severing the microtubule along its length. The protein katanin, a member of the AAA adenosine triphosphatase (ATPase) superfamily, uses ATP hydrolysis energy to sever and disassemble stable microtubules. Katanin may play a role in releasing microtubules from centrosomes, regulating assembly of the mitotic spindle, and accelerating microtubule turnover during cell cycle transitions (Hartman, J.J. and Vale, R.D. (1999) Science 286:782-785).

Microtubular aggregates are associated with several disorders. An extraskeletal myxoid chondrosarcoma tumor from human contained parallel arrays of microtubules within the rough endoplasmic reticulum (Suzuki, T. et al. (1988) J. Pathol. 156:51-57). Microtubular aggregates were also found in hepatocytes from chimpanzees infected with hepatitis C virus. Monoclonal antibodies prepared to these aggregates detect a protein called p44 (or microtubular aggregates protein) (Maeda, T. et al. (1989) J. Gen. Virol. 70:1401-1407). A human homolog of p44 is inducible by interferon- α and interferon- β , but not by interferon- γ . p44 protein may be a mediator in the antiviral action of interferon (Kitamura, A. et al. (1994) Eur. J. Biochem. 224:877-883).

Dynein-related Motor Proteins

Dyneins are (-) end-directed motor proteins which act on microtubules. Two classes of dyneins exist, cytosolic and axonemal. Cytosolic dyneins are responsible for translocation of materials along cytoplasmic microtubules; for example, transport from the nerve terminal to the cell body and transport of endocytic vesicles to lysosomes. Cytoplasmic dyneins are also reported to play a role in mitosis. Axonemal dyneins are responsible for the beating of flagella and cilia. Dynein on one microtubule doublet walks along the adjacent microtubule doublet. This sliding force produces bending forces that cause the flagellum or cilium to beat. Dyneins have a native mass between 1000 and 2000 kDa and contain either two or three force-producing heads driven by the hydrolysis of ATP. The heads are linked via stalks to a basal domain which is composed of a highly variable number of accessory intermediate and light chains.

Kinesin-related Motor Proteins

Kinesins are (+) end-directed motor proteins which act on microtubules. The prototypical kinesin molecule is involved in the transport of membrane-bound vesicles and organelles. This function is particularly important for axonal transport in neurons. Kinesin is also important in all cell types for the transport of vesicles from the Golgi complex to the endoplasmic reticulum. This role is critical for maintaining the identity and functionality of these secretory organelles.

Kinesins define a ubiquitous, conserved family of over 50 proteins that can be classified into at least 8 subfamilies based on primary amino acid sequence, domain structure, velocity of movement, and cellular function. (Reviewed in Moore, J.D. and S.A. Endow (1996) Bioessays 18:207-219; and Hoyt, A.M. (1994) Curr. Opin. Cell Biol. 6:63-68.) The prototypical kinesin molecule is a heterotetramer comprised of two heavy polypeptide chains (KHCs) and two light polypeptide chains (KLCs). The KHC subunits are typically referred to as "kinesin." KHC is about 1000 amino acids in length, and KLC is about 550 amino acids in length. Two KHCs dimerize to form a rod-shaped molecule with three distinct regions of secondary structure. At one end of the molecule is a globular motor domain that functions in ATP hydrolysis and microtubule binding. Kinesin motor domains are highly conserved and share over 70% identity. Beyond the motor domain is an α -helical coiled-coil region which mediates dimerization. At the other end of the molecule is a fan-shaped tail that associates with molecular cargo. The tail is formed by the interaction of the KHC C-termini with the two KLCs.

Members of the more divergent subfamilies of kinesins are called kinesin-related proteins (KRPs), many of which function during mitosis in eukaryotes (Hoyt, *supra*). Some KRPs are required for assembly of the mitotic spindle. *In vivo* and *in vitro* analyses suggest that these KRPs exert force on microtubules that comprise the mitotic spindle, resulting in the separation of spindle poles. Phosphorylation of KRP is required for this activity. Failure to assemble the mitotic spindle results in abortive mitosis and chromosomal aneuploidy, the latter condition being characteristic of cancer cells. In addition, a unique KRP, centromere protein E, localizes to the kinetochore of human mitotic chromosomes and may play a role in their segregation to opposite spindle poles.

Microfilaments and Associated Proteins

Actins

Microfilaments, cytoskeletal filaments with a diameter of 7-9 nm, are vital to cell locomotion, cell shape, cell adhesion, cell division, and muscle contraction. Assembly and disassembly of the microfilaments allow cells to change their morphology. Microfilaments are the polymerized form of actin, the most abundant intracellular protein in the eukaryotic cell. Human cells contain six isoforms of actin. The three α -actins are found in different kinds of muscle, nonmuscle β -actin and nonmuscle γ -actin are found in nonmuscle cells, and another γ -actin is found in intestinal smooth muscle cells. G-

actin, the monomeric form of actin, polymerizes into polarized, helical F-actin filaments, accompanied by the hydrolysis of ATP to ADP. Actin filaments associate to form bundles and networks, providing a framework to support the plasma membrane and determine cell shape. These bundles and networks are connected to the cell membrane. In muscle cells, thin filaments containing actin slide past thick
5 filaments containing the motor protein myosin during contraction. Other actin-related filaments are not part of the actin cytoskeleton, but rather associate with microtubules and dyenin.

Actin-Associated Proteins

Actin-associated proteins have roles in cross-linking, severing, and stabilization of actin filaments and in sequestering actin monomers. Several of the actin-associated proteins have multiple
10 functions. Bundles and networks of actin filaments are held together by actin cross-linking proteins. These proteins have two actin-binding sites, one for each filament. Short cross-linking proteins promote bundle formation while longer, more flexible cross-linking proteins promote network formation. Calmodulin-like calcium-binding domains in actin cross-linking proteins allow calcium regulation of cross-linking. Group I cross-linking proteins have unique actin-binding domains and include the 30 Kd
15 protein, EF-1a, fascin, and scruin. Group II cross-linking proteins have a 7,000-MW actin-binding domain and include villin and dematin. Group III cross-linking proteins have pairs of a 26,000-MW actin-binding domain and include alpha-actinin, fimbrin, spectrin, dystrophin, ABP 120, and filamin.

Severing proteins regulate the length of actin filaments by breaking them into short pieces or by blocking their ends. Severing proteins include gCAP39, severin (fragmin), gelsolin, and villin.
20 Capping proteins can cap the ends of actin filaments, but cannot break filaments. Capping proteins include CapZ, tropomodulin, and tensin.

Tensin, which is found in focal adhesions, also crosslinks actin filaments. Integrin activation by the extracellular matrix leads to the phosphorylation of tensin on tyrosine, serine, and threonine residues; this phosphorylation also occurs in cells transformed with oncogenes. Tensin has an SH2
25 domain and may bind to other tyrosine-phosphorylated proteins. (Lo, S.H. et al. (1997) J. Cell Biol. 136:1349-1361.) The N-terminus of tensin contains a region homologous to the catalytic domain of a putative tyrosine phosphatase (PTP) from *Saccharomyces cerevisiae*. This PTP domain in tensin may mediate binding interactions with phosphorylated polypeptides (Haynie, D.T. and Ponting, C.P. (1996) Protein Sci. 5:2643-2646). Mice which lack the tensin gene have kidney abnormalities, indicating that
30 the loss of tensin leads to weakening of focal adhesions in the kidney (Lo, *supra*).

The proteins thymosin and profilin sequester actin monomers in the cytosol, allowing a pool of unpolymerized actin to exist. Profilin may also stimulate F-actin formation by effectively lowering the critical concentration required for actin monomer addition (Gertler, F.B. et al. (1996) Cell 87:227-239).

The actin-associated proteins tropomyosin, troponin, and caldesmon regulate muscle

contraction in response to calcium. The tropomyosin proteins, found in muscle and nonmuscle cells, are α -helical and form coiled-coil dimers. Striated muscle tropomyosin mediates the interactions between the troponin complex and actin, regulating muscle contraction (PROSITE PDOC00290 Tropomyosins signature). The troponin complex is composed of troponin-T, troponin-I, and troponin-C. Troponin-T binds tropomyosin, linking troponin-I and troponin-C to tropomyosin.

The actin-associated proteins tropomyosin, troponin, and caldesmon regulate muscle contraction in response to calcium. The tropomyosin proteins, found in muscle and nonmuscle cells, are α -helical and form coiled-coil dimers. Striated muscle tropomyosin mediates the interactions between the troponin complex and actin, regulating muscle contraction. (PROSITE PDOC00290 Tropomyosins signature.) The troponin complex is composed of troponin-T, troponin-I, and troponin-C. Troponin-T binds tropomyosin, linking troponin-I and troponin-C to tropomyosin.

Many proteins involved in the regulation of actin assembly have characteristic protein-protein interaction domains, such as for example the calponin homology (CH) domain found in actin cross-linking proteins including alpha-actinin, spectrin, and fimbrin. Other proteins which are localized primarily in focal adhesions, macromolecular complexes which mediate the contact between extracellular matrix receptors and the cytoskeleton, contain protein-protein interaction motifs known as LIM domains. For example, zyxin is a protein that plays a role in the spatial control of actin assembly and contains three tandem LIM domains. Zyxin also interacts with alpha-actinin through its proline rich N-terminus (Beckerle, M. C. (1997) BioEssays 19:949-957).

Cytoskeletal proteins are implicated in several diseases. Pathologies such as muscular dystrophy, nephrotic syndrome, and dilated cardiomyopathy have been associated with differential expression of alpha-actinin-3 (Vainzof, M. et al. (1997) Neuropediatrics 28:223-228; Smoyer, W.E. and Mundel, P. (1998) J. Mol. Med. 76:172-183; and Sussman, M.A. et al. (1998) J. Clin. Invest. 101:51-61). Alpha-actinin and several MAPs are present in Hirano bodies, which are observed more frequently in the elderly and in patients with neurodegenerative diseases such as Alzheimer's disease (Maciver, S.K. and Harrington, C.R. (1995) Neuroreport. 6:1985-1988). Actinin-4, a novel actin-bundling protein, appears to be associated with the cell motility of metastatic cancer cells. Other disease associations include premature chromosome condensation which is frequently observed in dividing cells from tumor tissue (Murnane, J.P. (1995) Cancer Metastasis Rev. 14:17-29) and the significant roles of axonemal and assembly MAPs in viral pathogenesis (Sodeik, B. et al. (1997) J. Cell Biol. 136:1007-1021).

Intermediate Filaments and Associated Proteins

Intermediate filaments (IFs) are cytoskeletal fibers with a diameter of 10 nm, intermediate between that of microfilaments and microtubules. They serve structural roles in the cell, reinforcing cells and organizing cells into tissues. IFs are particularly abundant in epidermal cells and in neurons.

IFs are extremely stable, and, in contrast to microfilaments and microtubules, do not function in cell motility. IF proteins include acidic keratins, basic keratins, desmin, glial fibrillary acidic protein, vimentin, peripherin, neurofilaments, nestin, and lamins.

IFs have a central α -helical rod region interrupted by short nonhelical linker segments. The rod region is bracketed, in most cases, by non-helical head and tail domains. The rod regions of intermediate filament proteins associate to form a coiled-coil dimer. A highly ordered assembly process leads from the dimers to the IFs. Neither ATP nor GTP is needed for IF assembly, unlike that of microfilaments and microtubules.

IF-associated proteins (IFAPs) mediate the interactions of IFs with one another and with other cell structures. IFAPs cross-link IFs into a bundle, into a network, or to the plasma membrane, and may cross-link IFs to the microfilament and microtubule cytoskeleton. Microtubules and IFs in particular are closely associated. IFAPs include BPAG1, plakoglobin, desmoplakin I, desmoplakin II, plectin, ankyrin, filaggrin, and lamin B receptor.

The N-terminal portion of ankyrin consists of a repeated 33-amino acid motif, the ankyrin repeat, which is involved in specific protein-protein interactions. Variable regions within the motif are responsible for specific protein binding, such that different ankyrin repeats are involved in binding to tubulin, anion exchange protein, voltage-gated sodium channel, Na^+/K^+ -ATPase, and neurofascin. The ankyrin motif is also found in transcription factors, such as NF- κ -B, and in the yeast cell cycle proteins CDC10, SW14, and SW16. Proteins involved in tissue differentiation, such as *Drosophila* Notch and *C. elegans* LIN-12 and GLP-1, also contain ankyrin-like repeats. Lux et al. (1990; Nature 344:36-42) suggest that ankyrin-like repeats function as 'built-in' ankyrins and form binding sites for integral membrane proteins, tubulin, and other proteins.

Cytoskeletal-Membrane Anchors

Cytoskeletal fibers are attached to the plasma membrane by specific proteins. These attachments are important for maintaining cell shape and for muscle contraction. In erythrocytes, the spectrin-actin cytoskeleton is attached to cell membrane by three proteins, band 4.1, ankyrin, and adducin. Defects in this attachment result in abnormally shaped cells which are more rapidly degraded by the spleen, leading to anemia. In platelets, the spectrin-actin cytoskeleton is also linked to the membrane by ankyrin; a second actin network is anchored to the membrane by filamin. In muscle cells the protein dystrophin links actin filaments to the plasma membrane; mutations in the dystrophin gene lead to Duchenne muscular dystrophy. In adherens junctions and adhesion plaques the peripheral membrane proteins α -actinin and vinculin attach actin filaments to the cell membrane.

IFs are also attached to membranes by cytoskeletal-membrane anchors. The nuclear lamina is attached to the inner surface of the nuclear membrane by the lamin B receptor. Vimentin IFs are

attached to the plasma membrane by ankyrin and plectin. Desmosome and hemidesmosome membrane junctions hold together epithelial cells of organs and skin. These membrane junctions allow shear forces to be distributed across the entire epithelial cell layer, thus providing strength and rigidity to the epithelium. IFs in epithelial cells are attached to the desmosome by plakoglobin and desmoplakins.

- 5 The proteins that link IFs to hemidesmosomes are not known. Desmin IFs surround the sarcomere in muscle and are linked to the plasma membrane by paranemin, synemin, and ankyrin.

Proteins of the Erythrocyte Membrane Skeleton

- Distribution of oxygen throughout the vertebrate body is effected by red blood cells (erythrocytes). Oxygen diffuses from surrounding water or from the atmosphere through either gill
10 epithelium or pulmonary epithelial type I cells. Oxygen then diffuses through the blood capillary endothelium directly to the blood circulatory system and through the erythrocyte membrane and is stored as soluble oxyhemoglobin in the cytoplasm. Oxygen is released from hemoglobin at sites throughout the organism and diffuses out from the erythrocyte to other target cells. The structure of the erythrocyte membrane as well as that of other non-erythrocyte cells must be maintained to enable
15 efficient diffusion of oxygen to intracellular compartments.

- The erythrocyte membrane is comprised of i) a cholesterol-rich phospholipid bilayer in which many trans-bilayer proteins are embedded, ii) external glycosylphosphatidylinositol-anchored proteins (GPI-proteins), and iii) the erythrocyte or membrane skeleton that laminates the inner surface of the bilayer. The trans-bilayer proteins include anion exchangers, glycophorins, glucose transporters, and a
20 variety of cation transporters and pumps. The erythrocyte GPI-proteins include acetylcholinesterase and decay-accelerating factor (CD 55). The skeletal proteins are organized on the cortical, or cytoplasmic, face of the plasma membrane. These proteins include protein 4.1, protein p55, α - and β -spectrin, actin, and actin-binding proteins such as dematin, tropomyosin, and tropomodulin. α - and β -spectrin combine to form a heterotetramer in vivo. The spectrin heterotetramer organizes into a cortical
25 bidimensional network with a hexagonal mesh. The network is linked to trans-bilayer proteins through a protein complex comprising β -spectrin, ankyrin, anion exchanger, and protein 4.2 and through the "triangular" interaction between protein 4.1, glycophorin C, and protein p55. Structural and functional variants of erythrocyte membrane proteins have been found in a variety of tissues. Variants may be transcribed from multigene families, e.g., anion exchanger, ankyrin, or spectrin, or from single
30 gene families, e.g., protein 4.1 or protein 4.2. mRNA transcripts undergo tissue-specific alternative splicing. Many congenital hemolytic anemias result from mutations in the above-mentioned genes encoding erythrocyte membrane proteins. For example, hereditary elliptocytosis stems from an array of mutations in the spectrin genes at or near the head-to-head self-association region of the spectrin tetramer, or from mutations in the protein 4.1 gene which reduce levels of protein 4.1. In another

example, hereditary spherocytosis is associated with mutations in the ankyrin gene, the anion exchanger gene, the protein 4.2 gene, or the α - and β -spectrin genes. (Delaunay J. (1995) Transfus. Clin. Biol. 2:207-216.)

Protein 4.1 is an 80 kDa erythrocyte membrane protein with four functional domains. These domains include: i) a 30 kDa basic N-terminal domain, homologous to the ERM (Ezrin/Radixin/Moesin) family of actin- and transmembrane protein-binding proteins (Tsukita, S. et al. (1997) Trends Biochem. Sci. 22:53-58); ii) a 16 kDa hydrophilic domain containing a protein kinase C phosphorylation site; iii) a 10 kDa highly charged domain containing a cAMP-dependent protein kinase phosphorylation site critical for the interaction with spectrin and actin; and iv) a 22/24 kDa acidic domain. Protein 4.1 is a member of a structurally and functionally related protein 4.1 family. The protein 4.1 family is part of an evolutionarily related protein superfamily that includes many tyrosine phosphatases. (Baklouti, F. et al. (1997) Genomics 39:289-302.)

In contrast to the strictly cortical localization of protein 4.1 in mature enucleate erythrocytes, protein 4.1 epitopes have been observed throughout the cytoplasmic compartment and the nucleoskeleton in nucleated cells. In particular, protein 4.1 is present in the nucleoskeleton during interphase, in the mitotic spindle during mitosis, in perichromatin during telophase, and in the midbody during cytokinesis. (Krauss, S.W. et al. (1997) J. Cell Biol. 137:275-289.)

Differential expression of the protein 4.1 gene resulting in a number of mRNA splice variants has been observed in various human and rodent tissues. Comparison of the gene structure and mRNA splice variants revealed the extreme genomic sequence conservation of protein 4.1 between different species. The 5' UTR of both the human and rodent mRNA species has not been successfully identified and sequenced, possibly due to GC-rich regions therein which give rise to technical complications during nucleotide sequencing reactions. (Baklouti, *supra*; Conboy, J.G. (1988) Proc. Natl. Acad. Sci. 85:9062-9065.)

Analysis of proteins included in the ERM family of proteins has indicated that the N-terminal domain interacts with intracellular domains of transmembrane proteins such as CD44 and the C-terminal domain binds actin. Both interactions involve interactions with Rho-GTP protein complex, polyphosphoinositides, and serine/threonine kinase and tyrosine kinase activities. Many of the phosphorylation sites on ERM proteins are conserved. Although expression of ERM proteins *in vivo* is restricted to tissues such as endothelium, repression of ERM protein gene expression is released under conditions of cell culture. (Tsukita, *supra*.)

The cortical actin cytoskeleton participates in various membrane-based processes which necessitate a large amount of functional plasticity in the molecular components involved. A family of proteins homologous to band 4.1 is involved in the reorganization of the actin cytoskeleton in response

to various stimuli and probably plays a role in transmembrane signaling. This family includes tyrosine phosphatases, substrates of tyrosine kinases and a candidate for a tumor-suppressor gene. (Arpin M, et al. (1994) Curr. Opin. Cell Biol. 6:136-141.)

Disruptions in cytoskeletal protein interaction have been identified in a number of disease
 5 conditions or disorders. Neurofibromatosis type 2 is an autosomal dominant disease of the nervous system. Schwann cells isolated from patients with neurofibromatosis type 2 have characteristic morphology and growth parameters which differ from control Schwann cells. A gene associated with neurofibromatosis type 2 has been identified and is termed NF2. The NF2 gene product, known as schwannomin or merlin, is a member of the protein 4.1 superfamily, and mutations in the NF2 gene
 10 have been shown to be associated with the disease. (Rosenbaum, C. et al. (1998) Neurobiol. Dis. 5:55-64.) In addition, a form of psoriasis may be due to altered expression or distribution in epidermal epithelium of analogs of erythrocyte protein 4.1. (Shimizu, T. (1996) Histol. Histopathol. 11:495-501.) Erythrocytes carrying mutations in spectrin and protein 4.1 showed differing sensitivities to invasion by Plasmodium falciparum. (Facer, C.A. (1995) Parasitol Res. 81:52-57.) Furthermore, antibodies raised
 15 against erythrocyte protein 4.1 stained the majority of neurofibrillary tangles in the prefrontal cortex and hippocampus of brain tissue from patients with Alzheimer's disease. A 68 kDa protein was identified as the most likely brain analog of erythrocyte protein 4.1. (Sihag, R.K. et al. (1994) Brain Res. 656:14-26.)

The discovery of new cytoskeleton-associated proteins and the polynucleotides encoding them
 20 satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative, autoimmune/inflammatory, vesicle trafficking, neurological, cell motility, reproductive, and muscle disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of cytoskeleton-associated proteins.

25

SUMMARY OF THE INVENTION

The invention features purified polypeptides, cytoskeleton-associated proteins, referred to collectively as "CYSKP" and individually as "CYSKP-1," "CYSKP-2," "CYSKP-3," "CYSKP-4," "CYSKP-5," "CYSKP-6," "CYSKP-7," "CYSKP-8," "CYSKP-9," "CYSKP-10," "CYSKP-11,"
 30 "CYSKP-12," "CYSKP-13," "CYSKP-14," "CYSKP-15," "CYSKP-16," "CYSKP-17," "CYSKP-18," "CYSKP-19," "CYSKP-20," "CYSKP-21," "CYSKP-22," "CYSKP-23," "CYSKP-24," "CYSKP-25," "CYSKP-26," "CYSKP-27," "CYSKP-28," "CYSKP-29," "CYSKP-30," "CYSKP-31," "CYSKP-32," "CYSKP-33," and "CYSKP-34." In one aspect, the invention provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence

selected from the group consisting of SEQ ID NO:1-34, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, and d) an immunogenic
5 fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-34.

The invention further provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group
10 consisting of SEQ ID NO:1-34, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34. In one
15 alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-34. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:35-68.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting
20 of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, and d) an immunogenic fragment of a polypeptide having an amino
25 acid sequence selected from the group consisting of SEQ ID NO:1-34. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of
30 SEQ ID NO:1-34, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34. The method comprises a)

culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a

5 polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, and d) an immunogenic

10 fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34.

The invention further provides an isolated polynucleotide selected from the group consisting of

a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:35-68, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90%

15 identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:35-68, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample,

20 said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of

a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:35-68, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:35-68, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the

25 polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of

30 said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID

NO:35-68, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:35-68, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said
5 target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from
10 the group consisting of SEQ ID NO:1-34, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, and
15 a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-34. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional CYSKP, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an
20 agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, and
25 d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of
30 treating a disease or condition associated with decreased expression of functional CYSKP, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, b) a naturally occurring

polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional CYSKP, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in

altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:35-68, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

5 The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:35-68, ii) a
10 naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:35-68, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in
15 the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:35-68, ii) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:35-68, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the
20 polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated
25 biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the present invention.

30 Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog for polypeptides of the invention. The probability score for the match between each polypeptide and its GenBank homolog is also shown.

Table 3 shows structural features of polypeptide sequences of the invention, including predicted motifs and domains, along with the methods, algorithms, and searchable databases used for analysis of
35 the polypeptides.

Table 4 lists the cDNA and genomic DNA fragments which were used to assemble polynucleotide sequences of the invention, along with selected fragments of the polynucleotide sequences.

Table 5 shows the representative cDNA library for polynucleotides of the invention.

5 Table 6 provides an appendix which describes the tissues and vectors used for construction of the cDNA libraries shown in Table 5.

Table 7 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

10 DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will
15 be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so
20 forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described.
25 All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

30 "CYSKP" refers to the amino acid sequences of substantially purified CYSKP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of CYSKP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other

compound or composition which modulates the activity of CYSKP either by directly interacting with CYSKP or by acting on components of the biological pathway in which CYSKP participates.

An "allelic variant" is an alternative form of the gene encoding CYSKP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

10 "Altered" nucleic acid sequences encoding CYSKP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as CYSKP or a polypeptide with at least one functional characteristic of CYSKP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding CYSKP, and improper or unexpected hybridization to allelic variants, with 15 a locus other than the normal chromosomal locus for the polynucleotide sequence encoding CYSKP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent CYSKP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of CYSKP is retained. For example, negatively charged amino 20 acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and 25 valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence 30 to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of

CYSKP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of CYSKP either by directly interacting with CYSKP or by acting on components of the biological pathway in which CYSKP participates.

- 5 The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')₂, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind CYSKP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or
10 synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

- The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to
15 immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

- The term "antisense" refers to any composition capable of base-pairing with the "sense"
20 (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense
25 molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

- 30 The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic CYSKP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding CYSKP or fragments of CYSKP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Applied Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

| | Original Residue | Conservative Substitution |
|----|------------------|---------------------------|
| 25 | Ala | Gly, Ser |
| | Arg | His, Lys |
| | Asn | Asp, Gln, His |
| | Asp | Asn, Glu |
| | Cys | Ala, Ser |
| 30 | Gln | Asn, Glu, His |
| | Glu | Asp, Gln, His |
| | Gly | Ala |
| | His | Asn, Arg, Gln, Glu |
| | Ile | Leu, Val |
| 35 | Leu | Ile, Val |
| | Lys | Arg, Gln, Glu |
| | Met | Leu, Ile |
| | Phe | His, Met, Leu, Trp, Tyr |
| | Ser | Cys, Thr |

| | |
|-----|---------------|
| Thr | Ser, Val |
| Trp | Phe, Tyr |
| Tyr | His, Phe, Trp |
| Val | Ile, Leu, Thr |

5

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

10 A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one
15 biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

20 "Differential expression" refers to increased or upregulated; or decreased, downregulated, or absent gene or protein expression, determined by comparing at least two different samples. Such comparisons may be carried out between, for example, a treated and an untreated sample, or a diseased and a normal sample.

A "fragment" is a unique portion of CYSKP or the polynucleotide encoding CYSKP which is
25 identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid
30 residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present
35 embodiments.

A fragment of SEQ ID NO:35-68 comprises a region of unique polynucleotide sequence that

specifically identifies SEQ ID NO:35-68, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:35-68 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:35-68 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:35-68 and the region of SEQ ID NO:35-68 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-34 is encoded by a fragment of SEQ ID NO:35-68. A fragment of SEQ ID NO:1-34 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-34. For example, a fragment of SEQ ID NO:1-34 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-34. The precise length of a fragment of SEQ ID NO:1-34 and the region of SEQ ID NO:1-34 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "full length" polynucleotide sequence is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full length" polynucleotide sequence encodes a "full length" polypeptide sequence.

"Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search

Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions,

explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment
5 program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

10 Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

15 *Gap x drop-off: 50*

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for
20 example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to
25 describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid
30 sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity.

Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of
5 various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of CYSKP which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of
10 CYSKP which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

15 The term "modulate" refers to a change in the activity of CYSKP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of CYSKP.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or
20 synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding
25 sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs
30 preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an CYSKP may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by

cell type depending on the enzymatic milieu of CYSKP.

"Probe" refers to nucleic acid sequences encoding CYSKP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical
5 labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

10 Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the
15 specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel, F.M. et al. (1987) Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al. (1990) PCR
20 Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such
25 purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South
30 West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for

microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that
5 hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary
10 polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the
15 artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

20 Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs).
25 Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and
30 other moieties known in the art.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing CYSKP, nucleic acids encoding CYSKP, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

5 The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope
10 A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

 The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which
15 they are naturally associated.

 A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

 "Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers,
20 microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

 A "transcript image" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

 "Transformation" describes a process by which exogenous DNA is introduced into a recipient
25 cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed cells"
30 includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

 A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic

acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in
5 vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention
10 into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999)
15 set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant
20 identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternative splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic
25 variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at
30 least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a

certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human cytoskeleton-associated proteins (CYSKP), the polynucleotides encoding CYSKP, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, autoimmune/inflammatory, vesicle trafficking, neurological, cell motility, reproductive, and muscle disorders.

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO:) and an Incyte polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide sequence is denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown.

Table 2 shows sequences with homology to the polypeptides of the invention as identified by BLAST analysis against the GenBank protein (genpept) database. Columns 1 and 2 show the polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (Genbank ID NO:) of the nearest GenBank homolog. Column 4 shows the probability score for the match between each polypeptide and its GenBank homolog. Column 5 shows the annotation of the GenBank homolog along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Genetics Computer Group, Madison WI). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs. Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are cytoskeleton-associated proteins. For example, SEQ ID NO:31 is 34% identical to a *Caenorhabditis elegans* protein similar to mouse ankyrin

(GenBank ID g3879121) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 1.1×10^{-146} , which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:31 also contains Ank repeats as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. As a second example, SEQ ID NO:34 is 96% identical over 97 amino acids to human Intermediate Filament Associated Protein (GenBank ID 1333846) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 8.2×10^{-45} , which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. Data from BLAST analyses using the PRODOM database provide further corroborative evidence that SEQ ID NO:34 is a cytoskeleton protein. (See Table 3.) SEQ ID NO:1-30 and SEQ ID NO:32-33 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-34 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and the corresponding Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention. Column 3 shows the length of each polynucleotide sequence in basepairs. Column 4 lists fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:35-68 or that distinguish between SEQ ID NO:35-68 and related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages comprised of both cDNA and genomic DNA. These sequences were used to assemble the full length polynucleotide sequences of the invention. Columns 6 and 7 of Table 4 show the nucleotide start (5') and stop (3') positions of the cDNA and genomic sequences in column 5 relative to their respective full length sequences.

The identification numbers in Column 5 of Table 4 may refer specifically, for example, to Incyte cDNAs along with their corresponding cDNA libraries. For example, 3824958H1 is the identification number of an Incyte cDNA sequence, and BRAXNOT01 is the cDNA library from which it is derived. Incyte cDNAs for which cDNA libraries are not indicated were derived from pooled cDNA libraries (e.g., 71263527V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs (e.g., g2276318) which contributed to the assembly of the full length polynucleotide sequences. Alternatively, the identification numbers in column 5 may refer to coding regions predicted by Genscan analysis of genomic DNA. The Genscan-predicted coding sequences may

have been edited prior to assembly. (See Example IV.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. (See Example V.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon-stretching" algorithm. (See Example V.) In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in column 5 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

Table 5 shows the representative cDNA libraries for those full length polynucleotide sequences which were assembled using Incyte cDNA sequences. The representative cDNA library is the Incyte cDNA library which is most frequently represented by the Incyte cDNA sequences which were used to assemble and confirm the above polynucleotide sequences. The tissues and vectors which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

The invention also encompasses CYSKP variants. A preferred CYSKP variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the CYSKP amino acid sequence, and which contains at least one functional or structural characteristic of CYSKP.

The invention also encompasses polynucleotides which encode CYSKP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:35-68, which encodes CYSKP. The polynucleotide sequences of SEQ ID NO:35-68, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding CYSKP. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding CYSKP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:35-68 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:35-68. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of CYSKP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding CYSKP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the

invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring CYSKP, and all such variations are to be considered as being specifically disclosed.

5 Although nucleotide sequences which encode CYSKP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring CYSKP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding CYSKP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide
10 occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding CYSKP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

15 The invention also encompasses production of DNA sequences which encode CYSKP and CYSKP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding CYSKP or any fragment thereof.

20 Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:35-68 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in
25 "Definitions."

 Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or
30 combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing

system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding CYSKP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-

specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer
5 controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode CYSKP may be cloned in recombinant DNA molecules that direct expression of CYSKP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of
10 the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express CYSKP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter CYSKP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA
15 shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such
20 as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of CYSKP, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene
25 variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random
30 point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding CYSKP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) *Nucleic Acids Symp. Ser.* 7:215-223; and Horn, T. et al. (1980) *Nucleic Acids Symp. Ser.* 7:225-232.) Alternatively, CYSKP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide
5 synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) *Science* 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of CYSKP, or any part thereof, may be altered during direct synthesis and/or combined with sequences
10 from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) *Methods Enzymol.* 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing.
15 (See, e.g., Creighton, supra, pp. 28-53.)

In order to express a biologically active CYSKP, the nucleotide sequences encoding CYSKP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and
20 inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding CYSKP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding CYSKP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding CYSKP and its initiation codon and upstream regulatory sequences are inserted
25 into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers
30 appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding CYSKP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in

vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences
5 encoding CYSKP. These include, but are not limited to, microorganisms such as bacteria transformed
with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with
yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus);
plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or
tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or
10 animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster
(1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA
91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO
J. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New
York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and
15 Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses,
adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for
delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See; e.g., Di
Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci.
USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al.
20 (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.)
The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending
upon the use intended for polynucleotide sequences encoding CYSKP. For example, routine cloning,
subcloning, and propagation of polynucleotide sequences encoding CYSKP can be achieved using a
25 multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPT1 plasmid
(Life Technologies). Ligation of sequences encoding CYSKP into the vector's multiple cloning site
disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed
bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro
transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested
30 deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem.
264:5503-5509.) When large quantities of CYSKP are needed, e.g. for the production of antibodies,
vectors which direct high level expression of CYSKP may be used. For example, vectors containing the
strong, inducible SP6 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of CYSKP. A number of vectors

containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra;
5 Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184.)

Plant systems may also be used for expression of CYSKP. Transcription of sequences encoding CYSKP may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.*
10 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill,
15 New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding CYSKP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain
20 infective virus which expresses CYSKP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of
25 DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of CYSKP in cell lines is preferred. For example, sequences encoding CYSKP can be transformed into
30 cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the

introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk⁻* and *apr⁻* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; 5 Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) 10 Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used. 15 These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the 20 sequence encoding CYSKP is inserted within a marker gene sequence, transformed cells containing sequences encoding CYSKP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding CYSKP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

25 In general, host cells that contain the nucleic acid sequence encoding CYSKP and that express CYSKP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

30 Immunological methods for detecting and measuring the expression of CYSKP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on CYSKP is preferred, but a competitive binding

assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

5 A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding CYSKP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding CYSKP, or any fragments thereof, may be cloned into a vector
10 for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease
15 of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding CYSKP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence
20 and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode CYSKP may be designed to contain signal sequences which direct secretion of CYSKP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the
25 polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture
30 Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding CYSKP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric CYSKP protein

containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of CYSKP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the CYSKP encoding sequence and the heterologous protein sequence, so that CYSKP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled CYSKP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ³⁵S-methionine.

CYSKP of the present invention or fragments thereof may be used to screen for compounds that specifically bind to CYSKP. At least one and up to a plurality of test compounds may be screened for specific binding to CYSKP. Examples of test compounds include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of CYSKP, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, e.g., Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which CYSKP binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express CYSKP, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing CYSKP or cell membrane fractions which contain CYSKP are then contacted with a test compound and binding, stimulation, or inhibition of activity of either CYSKP or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with CYSKP, either in solution or affixed to a solid support, and detecting the binding of CYSKP to the compound.

- 5 Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

- 10 CYSKP of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of CYSKP. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for CYSKP activity, wherein CYSKP is combined with at least one test compound, and the activity of CYSKP in the presence of a test compound is compared with the activity of CYSKP in the absence of the test compound. A change in the activity of CYSKP in the presence of the test compound is
15 indicative of a compound that modulates the activity of CYSKP. Alternatively, a test compound is combined with an in vitro or cell-free system comprising CYSKP under conditions suitable for CYSKP activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of CYSKP may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

- 20 In another embodiment, polynucleotides encoding CYSKP or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent Number 5,175,383 and U.S. Patent Number 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo
25 and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D.
30 (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding CYSKP may also be manipulated in vitro in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998)

5 Science 282:1145-1147).

Polynucleotides encoding CYSKP can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding CYSKP is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae
10 are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress CYSKP, e.g., by secreting CYSKP in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

THERAPEUTICS

15 Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of CYSKP and cytoskeleton-associated proteins. In addition, the expression of CYSKP is closely associated with lung, reproductive (including placenta), neural (including brain), adrenal, endothelial, kidney, and spleen tissue, as well as with ovarian, breast, and testicular tumor tissue. Therefore, CYSKP appears to play a role in cell proliferative, autoimmune/inflammatory,
20 vesicle trafficking, neurological, cell motility, reproductive, and muscle disorders. In the treatment of disorders associated with increased CYSKP expression or activity, it is desirable to decrease the expression or activity of CYSKP. In the treatment of disorders associated with decreased CYSKP expression or activity, it is desirable to increase the expression or activity of CYSKP.

Therefore, in one embodiment, CYSKP or a fragment or derivative thereof may be
25 administered to a subject to treat or prevent a disorder associated with decreased expression or activity of CYSKP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia,
30 lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory

- distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with
- 5 lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus,
- 10 systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a vesicle trafficking disorder such as cystic fibrosis, glucose-galactose malabsorption syndrome, hypercholesterolemia, diabetes mellitus, diabetes insipidus, hyper- and hypoglycemia, Grave's disease, goiter, Cushing's disease, and Addison's disease,
- 15 gastrointestinal disorders including ulcerative colitis, gastric and duodenal ulcers, other conditions associated with abnormal vesicle trafficking, including acquired immunodeficiency syndrome (AIDS), allergies including hay fever, asthma, and urticaria (hives), autoimmune hemolytic anemia, proliferative glomerulonephritis, inflammatory bowel disease, multiple sclerosis, myasthenia gravis, rheumatoid and osteoarthritis, scleroderma, Chediak-Higashi and Sjogren's syndromes, systemic lupus erythematosus,
- 20 toxic shock syndrome, traumatic tissue damage, and viral, bacterial, fungal, helminthic, and protozoal infections; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and
- 25 other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal
- 30 syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety,

and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a cell motility disorder such as ankylosing spondylitis, Chediak-Higashi syndrome, Duchenne
5 and Becker muscular dystrophy, intrahepatic cholestasis, myocardial hyperplasia, cardiomyopathy, early onset peridontitis, cancers such as adenocarcinoma, ovarian carcinoma, and chronic myelogenous leukemia, and bacterial and helminthic infections; a reproductive disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, endometriosis, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian
10 hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, ectopic pregnancy, teratogenesis, cancer of the breast, fibrocystic breast disease, galactorrhea, a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, gynecomastia, hypergonadotropic and hypogonadotropic hypogonadism, pseudohermaphroditism,
15 azoospermia, premature ovarian failure, acrosin deficiency, delayed puberty, retrograde ejaculation and anejaculation, haemangioblastomas, cystsphaeochromocytomas, paraganglioma, cystadenomas of the epididymis, and endolymphatic sac tumours; and a muscle disorder such as myocarditis, Duchenne's muscular dystrophy, Becker's muscular dystrophy, myotonic dystrophy, central core disease, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, infectious myositis,
20 polymyositis, dermatomyositis, inclusion body myositis, thyrotoxic myopathy, and ethanol myopathy, angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, and pheochromocytoma, and myopathies including cardiomyopathy, encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, and ophthalmoplegia.

25 In another embodiment, a vector capable of expressing CYSKP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of CYSKP including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified CYSKP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a
30 disorder associated with decreased expression or activity of CYSKP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of CYSKP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of CYSKP including, but not limited to, those listed above.

In a further embodiment, an antagonist of CYSKP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of CYSKP. Examples of such disorders include, but are not limited to, those cell proliferative, autoimmune/inflammatory, vesicle trafficking, neurological, cell motility, reproductive, and muscle disorders described above. In one
5 aspect, an antibody which specifically binds CYSKP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express CYSKP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding CYSKP may be administered to a subject to treat or prevent a disorder associated with
10 increased expression or activity of CYSKP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination
15 of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of CYSKP may be produced using methods which are generally known in the art. In particular, purified CYSKP may be used to produce antibodies or to screen libraries of
20 pharmaceutical agents to identify those which specifically bind CYSKP. Antibodies to CYSKP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with CYSKP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols,
30 polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to CYSKP have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or

fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of CYSKP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to CYSKP may be prepared using any technique which provides for the
5 production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) *Nature* 256:495-497; Kozbor, D. et al. (1985) *J. Immunol. Methods* 81:31-42; Cote, R.J. et al. (1983) *Proc. Natl. Acad. Sci. USA* 80:2026-2030; and Cole, S.P. et al. (1984) *Mol. Cell Biol.* 62:109-120.)

10 In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) *Proc. Natl. Acad. Sci. USA* 81:6851-6855; Neuberger, M.S. et al. (1984) *Nature* 312:604-608; and Takeda, S. et al. (1985) *Nature* 314:452-454.) Alternatively, techniques described for the production of single
15 chain antibodies may be adapted, using methods known in the art, to produce CYSKP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) *Proc. Natl. Acad. Sci. USA* 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population
20 or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:3833-3837; Winter, G. et al. (1991) *Nature* 349:293-299.)

Antibody fragments which contain specific binding sites for CYSKP may also be generated. For example, such fragments include, but are not limited to, F(ab')₂ fragments produced by pepsin
25 digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')₂ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) *Science* 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired
30 specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between CYSKP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering CYSKP epitopes is generally used, but a competitive binding assay may also be

employed (Pound, *supra*).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for CYSKP. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of CYSKP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple CYSKP epitopes, represents the average affinity, or avidity, of the antibodies for CYSKP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular CYSKP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the CYSKP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of CYSKP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of CYSKP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, *supra*, and Coligan et al. *supra*.)

In another embodiment of the invention, the polynucleotides encoding CYSKP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding CYSKP. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding CYSKP. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J. et al. (1995)

9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood* 76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Boado, R.J. et al. (1998) *J. Pharm. Sci.* 87(11):1308-1315; and Morris, M.C. et al. (1997) *Nucleic Acids Res.* 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding CYSKP may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassamias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Somia (1997) *Nature* 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) *Nature* 335:395-396; Poeschla, E. et al. (1996) *Proc. Natl. Acad. Sci. USA.* 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in CYSKP expression or regulation causes disease, the expression of CYSKP from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in CYSKP are treated by constructing mammalian expression vectors encoding CYSKP and introducing these vectors by mechanical means into CYSKP-deficient cells. Mechanical transfer technologies for use with cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) *Annu. Rev. Biochem.* 62:191-217; Ivics, Z. (1997) *Cell* 91:501-510; Boulay, J-L. and H. Récipon (1998) *Curr. Opin. Biotechnol.* 9:445-450).

Expression vectors that may be effective for the expression of CYSKP include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). CYSKP may be expressed

5 using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) *Proc. Natl. Acad. Sci. USA* 89:5547-5551; Gossen, M. et al. (1995) *Science* 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) *Curr. Opin. Biotechnol.* 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the

10 ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and Blau, H.M. *supra*), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding CYSKP from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID

15 TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) *Virology* 52:456-467), or by electroporation (Neumann, E. et al. (1982) *EMBO J.* 1:841-845). The introduction of DNA to primary cells requires modification of these

20 standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to CYSKP expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding CYSKP under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive

25 element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) *Proc. Natl. Acad. Sci. USA* 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for

30 receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) *J. Virol.* 61:1647-1650; Bender, M.A. et al. (1987) *J. Virol.* 61:1639-1646; Adam, M.A. and A.D. Miller (1988) *J. Virol.* 62:3802-3806; Dull, T. et al. (1998) *J. Virol.* 72:8463-8471; Zufferey, R. et al. (1998) *J. Virol.* 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a

method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference.

Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4⁺ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) *J. Virol.* 71:7020-7029; Bauer, G. et al. (1997) *Blood* 89:2259-2267; Bonyhadi, M.L. (1997) *J. Virol.* 71:4707-4716; Ranga, U. et al. (1998) *Proc. Natl. Acad. Sci. USA* 95:1201-1206; Su, L. (1997) *Blood* 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding CYSKP to cells which have one or more genetic abnormalities with respect to the expression of CYSKP. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) *Transplantation* 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) *Annu. Rev. Nutr.* 19:511-544 and Verma, I.M. and N. Somia (1997) *Nature* 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding CYSKP to target cells which have one or more genetic abnormalities with respect to the expression of CYSKP. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing CYSKP to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) *Exp. Eye Res.* 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary

skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding CYSKP to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotechnol. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for CYSKP into the alphavirus genome in place of the capsid-coding region results in the production of a large number of CYSKP-coding RNAs and the synthesis of high levels of CYSKP in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of CYSKP into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding CYSKP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by

scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis.

10 Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding CYSKP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

15 RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and

20 wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding CYSKP. Compounds which may be effective in altering expression of a specific polynucleotide may include,

25 but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased

30 CYSKP expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding CYSKP may be therapeutically useful, and in the treatment of disorders associated with decreased CYSKP expression or activity, a compound which specifically promotes expression of the polynucleotide encoding CYSKP may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in

altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound
5 based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding CYSKP is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an in vitro cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding CYSKP are assayed
10 by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding CYSKP. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide
15 exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a Schizosaccharomyces pombe gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruce, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruce, T.W. et al. (2000) U.S. Patent No. 6,022,691).

25 Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a composition which
35 generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient.

Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of CYSKP, antibodies to CYSKP, and mimetics, agonists, antagonists, or inhibitors of CYSKP.

- 5 The compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Compositions for pulmonary administration may be prepared in liquid or dry powder form.

- 10 These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

- 20 Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising CYSKP or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, CYSKP or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

- 25 For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example CYSKP or fragments thereof, antibodies of CYSKP, and agonists, antagonists or inhibitors of CYSKP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by

standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Compositions which exhibit large
5 therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

10 The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy.
15 Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg , up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art.

20 Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind CYSKP may be used for the
25 diagnosis of disorders characterized by expression of CYSKP, or in assays to monitor patients being treated with CYSKP or agonists, antagonists, or inhibitors of CYSKP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for CYSKP include methods which utilize the antibody and a label to detect CYSKP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and
30 may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring CYSKP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of CYSKP expression. Normal or standard values for CYSKP expression are established by combining body fluids or cell extracts

taken from normal mammalian subjects, for example, human subjects, with antibodies to CYSKP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of CYSKP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values.

- 5 Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding CYSKP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of CYSKP may be correlated with
10 disease. The diagnostic assay may be used to determine absence, presence, and excess expression of CYSKP, and to monitor regulation of CYSKP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding CYSKP or closely related molecules may be used to identify nucleic acid sequences which encode CYSKP. The specificity of the probe, whether it is made
15 from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding CYSKP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50%
20 sequence identity to any of the CYSKP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:35-68 or from genomic sequences including promoters, enhancers, and introns of the CYSKP gene.

Means for producing specific hybridization probes for DNAs encoding CYSKP include the cloning of polynucleotide sequences encoding CYSKP or CYSKP derivatives into vectors for the
25 production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ³²P or ³⁵S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

30 Polynucleotide sequences encoding CYSKP may be used for the diagnosis of disorders associated with expression of CYSKP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including

adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with

5 lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus,

10 systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a vesicle trafficking disorder such as cystic fibrosis, glucose-galactose malabsorption syndrome, hypercholesterolemia, diabetes mellitus, diabetes insipidus, hyper- and hypoglycemia, Grave's disease, goiter, Cushing's disease, and Addison's disease,

20 gastrointestinal disorders including ulcerative colitis, gastric and duodenal ulcers, other conditions associated with abnormal vesicle trafficking, including acquired immunodeficiency syndrome (AIDS), allergies including hay fever, asthma, and urticaria (hives), autoimmune hemolytic anemia, proliferative glomerulonephritis, inflammatory bowel disease, multiple sclerosis, myasthenia gravis, rheumatoid and osteoarthritis, scleroderma, Chediak-Higashi and Sjogren's syndromes, systemic lupus erythematosus,

25 toxic shock syndrome, traumatic tissue damage, and viral, bacterial, fungal, helminthic, and protozoal infections; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and

30 other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal

syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a cell motility disorder such as ankylosing spondylitis, Chediak-Higashi syndrome, Duchenne and Becker muscular dystrophy, intrahepatic cholestasis, myocardial hyperplasia, cardiomyopathy, early onset periodontitis, cancers such as adenocarcinoma, ovarian carcinoma, and chronic myelogenous leukemia, and bacterial and helminthic infections; a reproductive disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, endometriosis, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, ectopic pregnancy, teratogenesis, cancer of the breast, fibrocystic breast disease, galactorrhea, a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, gynecomastia, hypergonadotropic and hypogonadotropic hypogonadism, pseudohermaphroditism, azoospermia, premature ovarian failure, acrosin deficiency, delayed puberty, retrograde ejaculation and anejaculation, haemangioblastomas, cystsphaeochromocytomas, paraganglioma, cystadenomas of the epididymis, and endolymphatic sac tumours; and a muscle disorder such as myocarditis, Duchenne's muscular dystrophy, Becker's muscular dystrophy, myotonic dystrophy, central core disease, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, infectious myositis, polymyositis, dermatomyositis, inclusion body myositis, thyrotoxic myopathy, and ethanol myopathy, angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, and pheochromocytoma, and myopathies including cardiomyopathy, encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, and ophthalmoplegia. The polynucleotide sequences encoding CYSKP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered CYSKP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding CYSKP may be useful in assays that

detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding CYSKP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding CYSKP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

10 In order to provide a basis for the diagnosis of a disorder associated with expression of CYSKP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding CYSKP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

30 Additional diagnostic uses for oligonucleotides designed from the sequences encoding CYSKP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding CYSKP, or a fragment of a polynucleotide complementary to the polynucleotide encoding CYSKP, and will be employed under optimized conditions for identification of a specific gene or condition.

Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding CYSKP may be used to detect single nucleotide polymorphisms (SNPs). SNPs are
5 substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding CYSKP are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal
10 tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed *in silico* SNP
15 (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high
20 throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of CYSKP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) *J. Immunol. Methods* 159:235-244; Duplaa, C. et al. (1993) *Anal. Biochem.* 212:229-236.) The speed of quantitation of multiple samples may be
25 accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray
30 can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the

activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

5 In another embodiment, CYSKP, fragments of CYSKP, or antibodies specific for CYSKP may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of
10 gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of
15 transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies,
20 or other biological samples. The transcript image may thus reflect gene expression in vivo, as in the case of a tissue or biopsy sample, or in vitro, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental
25 compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties.
30 These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data

after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released 5 February 29, 2000, available at <http://www.niehs.nih.gov/oc/news/toxchip.htm>.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated 10 biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

15 Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given 20 conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins are 25 visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein 30 spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for CYSKP to quantify the levels of CYSKP expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) *Anal. Biochem.* 270:103-111; 5 Mendoze, L.G. et al. (1999) *Biotechniques* 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation 10 between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) *Electrophoresis* 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

15 In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound 20 in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized 25 by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) *Proc. Natl. Acad. Sci.* 30 USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) *Proc. Natl. Acad. Sci.* USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding CYSKP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, for example, Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding CYSKP on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, CYSKP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a

solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between CYSKP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with CYSKP, or fragments thereof, and washed. Bound CYSKP is then detected by methods well known in the art. Purified CYSKP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding CYSKP specifically compete with a test compound for binding CYSKP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with CYSKP.

In additional embodiments, the nucleotide sequences which encode CYSKP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. 60/201,960, U.S. Ser. No. 60/202,729, U.S. Ser. No. 60/209,705, U.S. Ser. No. 60/210,149, and U.S. Ser. No. 60/213,215, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA) and shown in Table 4, column 5. Some tissues were homogenized

and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)+ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERScript plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), or pINCY (Incyte Genomics, Palo Alto CA), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) *Anal. Biochem.* 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using
5 PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation
10 such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).
15 Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997,
20 supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA
25 sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and hidden Markov model (HMM)-based protein family databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary structures of gene families. See, for example, Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365.) The queries were
30 performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages

were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide of the invention may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently
5 analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are
10 generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold
15 parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity
20 between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:35-68. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 4.

25 IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative cytoskeleton-associated proteins were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpr1 and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and S. Karlin (1997) J. Mol. Biol. 268:78-94, and Burge, C. and S. Karlin
30 (1998) Curr. Opin. Struct. Biol. 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode cytoskeleton-associated proteins, the encoded polypeptides were analyzed by querying

against PFAM models for cytoskeleton-associated proteins. Potential cytoskeleton-associated proteins were also identified by homology to Incyte cDNA sequences that had been annotated as cytoskeleton-associated proteins. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpr public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

15 "Stitched" Sequences

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpr public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

"Stretched" Sequences

Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST analysis to either Incyte cDNA sequences or GenScan exon predicted sequences described in Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs (HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions may occur in the chimeric protein with respect to the original GenBank protein homolog. The GenBank protein homolog, the chimeric protein, or both were used as probes to search for homologous genomic sequences from the public human genome databases. Partial DNA sequences were therefore "stretched" or extended by the addition of homologous genomic sequences. The resultant stretched sequences were examined to determine whether it contained a complete gene.

VI. Chromosomal Mapping of CYSKP Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:35-68 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:35-68 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

In this manner, SEQ ID NO:44 was mapped to chromosome 17 within the interval from

62.90 to 64.20 centiMorgans, SEQ ID NO:49 was mapped to chromosome 14 within the interval from 73.70 to 76.40 centiMorgans, SEQ ID NO:50 was mapped to chromosome 8 within the interval from 25.80 to 40.30 centiMorgans, SEQ ID NO:54 was mapped to chromosome 1 within the interval from 117.6 to 132.4 centiMorgans, SEQ ID NO:64 was mapped to chromosome 4 within the interval from 56.7 to 60.5 centiMorgans, and SEQ ID NO:65 was mapped to chromosome 5 within the interval from 141.40 to 142.60 centiMorgans.

VII. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel (1995) supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum} \{ \text{length}(\text{Seq. 1}), \text{length}(\text{Seq. 2}) \}}$$

20

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

Alternatively, polynucleotide sequences encoding CYSKP are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled,

at least in part, with overlapping Incyte cDNA sequences (see Example III). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding CYSKP. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

VIII. Extension of CYSKP Encoding Polynucleotides

Full length polynucleotide sequences were also produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+} , $(NH_4)_2SO_4$, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μ l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II
5 (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μ l to 10 μ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and
10 sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs,
15 and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1:
20 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC
25 DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, full length polynucleotide sequences are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

30 IX. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:35-68 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National

Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of [γ - 32 P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

X. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, *supra.*), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Skena (1999), *supra.*). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Skena, M. et al. (1995) *Science* 270:467-470; Shalon, D. et al. (1996) *Genome Res.* 6:639-645; Marshall, A. and J. Hodgson (1998) *Nat. Biotechnol.* 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorption and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on

the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/ μ l oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/ μ l RNase inhibitor, 500 μ M dATP, 500 μ M dGTP, 500 μ M dTTP, 40 μ M dCTP, 40 μ M dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)⁺ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)⁺ RNAs are synthesized by in vitro transcription from non-coding yeast genomic DNA. After incubation at 37° C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85° C to stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μ l 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 μ g. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110° C oven.

Array elements are applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522, incorporated herein by reference. 1 μ l of the array element DNA, at an average concentration of 100 ng/ μ l, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60° C followed by washes in
5 0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 µl of sample mixture consisting of 0.2 µg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65° C for 5 minutes and is aliquoted onto the microarray surface and covered with
10 an 1.8 cm² coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 µl of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60° C. The arrays are washed for 10 min at 45° C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45° C in a second wash buffer (0.1X SSC), and dried.

Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide
20 containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477,
25 Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different
35 fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially

expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC
5 computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

10 A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

XI. Complementary Polynucleotides

15 Sequences complementary to the CYSKP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring CYSKP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of CYSKP. To inhibit transcription, a
20 complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the CYSKP-encoding transcript.

XII. Expression of CYSKP

Expression and purification of CYSKP is achieved using bacterial or virus-based expression
25 systems. For expression of CYSKP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac (tac)* hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3).

30 Antibiotic resistant bacteria express CYSKP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of CYSKP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding CYSKP by either homologous recombination or bacterial-mediated

transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, CYSKP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from CYSKP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified CYSKP obtained by these methods can be used directly in the assays shown in Examples XVI and XVII, where applicable.

XIII. Functional Assays

CYSKP function is assessed by expressing the sequences encoding CYSKP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT (Life Technologies) and PCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in

cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated
5 Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of CYSKP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding CYSKP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human
10 immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding CYSKP and other genes of interest can be analyzed by northern analysis or microarray techniques.

15 XIV. Production of CYSKP Specific Antibodies

CYSKP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the CYSKP amino acid sequence is analyzed using LASERGENE software
20 (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A
25 peptide synthesizer (Applied Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-CYSKP activity by, for example, binding the peptide or CYSKP to a substrate, blocking with
30 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XV. Purification of Naturally Occurring CYSKP Using Specific Antibodies

Naturally occurring or recombinant CYSKP is substantially purified by immunoaffinity chromatography using antibodies specific for CYSKP. An immunoaffinity column is constructed by covalently coupling anti-CYSKP antibody to an activated chromatographic resin, such as

CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing CYSKP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of CYSKP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/CYSKP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and CYSKP is collected.

XVI. Identification of Molecules Which Interact with CYSKP

CYSKP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled CYSKP, washed, and any wells with labeled CYSKP complex are assayed. Data obtained using different concentrations of CYSKP are used to calculate values for the number, affinity, and association of CYSKP with the candidate molecules.

Alternatively, molecules interacting with CYSKP are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989) *Nature* 340:245-246, or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

CYSKP may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

XVII. Demonstration of CYSKP Activity

A microtubule motility assay for CYSKP measures motor protein activity. In this assay, recombinant CYSKP is immobilized onto a glass slide or similar substrate. Taxol-stabilized bovine brain microtubules (commercially available) in a solution containing ATP and cytosolic extract are perfused onto the slide. Movement of microtubules as driven by CYSKP motor activity can be visualized and quantified using video-enhanced light microscopy and image analysis techniques. CYSKP activity is directly proportional to the frequency and velocity of microtubule movement.

Alternatively, an assay for CYSKP measures the formation of protein filaments in vitro. A solution of CYSKP at a concentration greater than the "critical concentration" for polymer assembly is applied to carbon-coated grids. Appropriate nucleation sites may be supplied in the solution. The grids are negative stained with 0.7% (w/v) aqueous uranyl acetate and examined by electron microscopy. The appearance of filaments of approximately 25 nm (microtubules), 8 nm (actin), or 10 nm (intermediate filaments) is a demonstration of protein activity.

Alternatively, an assay for CYSKP measures the binding affinity of CYSKP for actin as described by Hammell, R.L. and Hitchcock-DeGregori, S.E. (1997, J. Biol. Chem. 272:22409-22416). CYSKP and actin are prepared from in vitro recombinant cDNA expression systems and the N-terminus of CYSKP is acetylated using methods well known in the art. Binding of N-terminal acetyl-
5 CYSKP to actin is measured by cosedimentation at 25 °C in a Beckman model TL-100 centrifuge as described. The bound and free CYSKP are determined by quantitative densitometry of SDS-polyacrylamide gels stained with Coomassie Blue. Apparent binding constants (K_{app}) and Hill coefficients (H) are determined by using methods well known in the art to fit the data to the equation as described by Hammell and Hitchcock-DeGregori (1997, supra). The CYSKP:actin ratio, determined
10 using densitometry, is normalized. Hammell and Hitchcock-DeGregori (1997, supra) have shown that saturation of binding corresponds to a CYSKP:actin molar ratio of 0.14, a stoichiometry of 1 CYSKP:7 actin. The binding of CYSKP to actin is proportional to the CYSKP activity.

Alternatively, CYSKP activity is measured as ability to bind to microtubules. Microtubules are purified from adult rat brain by reversible assembly (Vallee, R. B. (1982) Methods Enzymol.
15 134:89-104) or the taxol method (Vallee, R. B. (1982) J. Cell Biol. 92:435-442) using PEM buffer (100 mM PIPES, pH 6.6, 1mM EGTA, 1mM $MgSO_4$). To separate the MAPs from tubulin, the pellets from twice-cycled microtubules are resuspended in PEM buffer and applied to a 0.1 M $MgSO_4$ -saturated phosphocellulose column as described by Sloboda, R. D. and Rosenbaum, J. L. ((1982) Methods Enzymol. 85:409-416). The fractions containing protein are applied to a second
20 phosphocellulose column. In a total volume of 100 ml, 20 ml of CYSKP (250 mg/ml) is added to 80 ml of whole microtubules (450 mg/ml) or tubulin (300 mg/ml) and incubated at 37 °C for 10 minutes in the presence of 1 mM GTP and 50 mM taxol. The suspension is centrifuged, the supernatant is removed, and the microtubule pellet is resuspended to the original reaction volume in PEM buffer. To assess the partitioning of CYSKP between the supernatant and pellet fractions, equal amounts of supernatant and
25 resuspended pellet are placed in SDS sample buffer and assayed on a 5-20% gradient SDS polyacrylamide gel stained with Coomassie Brilliant Blue. The amount of CYSKP in the pellet fraction is proportional to the binding of CYSKP to microtubules.

Alternatively, CYSKP activity is associated with its ability to form protein-protein complexes and is measured by its ability to regulate growth characteristics of NIH3T3 mouse fibroblast cells. A
30 cDNA encoding CYSKP is subcloned into an appropriate eukaryotic expression vector. This vector is transfected into NIH3T3 cells using methods known in the art. Transfected cells are compared with non-transfected cells for the following quantifiable properties: growth in culture to high density, reduced attachment of cells to the substrate, altered cell morphology, and ability to induce tumors when injected into immunodeficient mice. The activity of CYSKP is proportional to the extent of increased growth or

frequency of altered cell morphology in NIH3T3 cells transfected with CYSKP.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention.

- 5 Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

| Incyte Project ID | Polypeptide SEQ ID NO: | Incyte Polypeptide ID | Polynucleotide SEQ ID NO: | Incyte Polynucleotide ID |
|----------------------|---------------------------|--------------------------|------------------------------|-----------------------------|
| 1889577 | 1 | 1889577CD1 | 35 | 1889577CB1 |
| 2427982 | 2 | 2427982CD1 | 36 | 2427982CB1 |
| 2470833 | 3 | 2470833CD1 | 37 | 2470833CB1 |
| 2080579 | 4 | 2080579CD1 | 38 | 2080579CB1 |
| 2156553 | 5 | 2156553CD1 | 39 | 2156553CB1 |
| 2182855 | 6 | 2182855CD1 | 40 | 2182855CB1 |
| 2242106 | 7 | 2242106CD1 | 41 | 2242106CB1 |
| 2726877 | 8 | 2726877CD1 | 42 | 2726877CB1 |
| 2738233 | 9 | 2738233CD1 | 43 | 2738233CB1 |
| 1833116 | 10 | 1833116CD1 | 44 | 1833116CB1 |
| 001799 | 11 | 001799CD1 | 45 | 001799CB1 |
| 119814 | 12 | 119814CD1 | 46 | 119814CB1 |
| 1295420 | 13 | 1295420CD1 | 47 | 1295420CB1 |
| 1309364 | 14 | 1309364CD1 | 48 | 1309364CB1 |
| 1315267 | 15 | 1315267CD1 | 49 | 1315267CB1 |
| 1403289 | 16 | 1403289CD1 | 50 | 1403289CB1 |
| 1607607 | 17 | 1607607CD1 | 51 | 1607607CB1 |
| 1660025 | 18 | 1660025CD1 | 52 | 1660025CB1 |
| 1796836 | 19 | 1796836CD1 | 53 | 1796836CB1 |
| 2880670 | 20 | 2880670CD1 | 54 | 2880670CB1 |
| 2913976 | 21 | 2913976CD1 | 55 | 2913976CB1 |
| 3092084 | 22 | 3092084CD1 | 56 | 3092084CB1 |
| 3882482 | 23 | 3882482CD1 | 57 | 3882482CB1 |
| 4933451 | 24 | 4933451CD1 | 58 | 4933451CB1 |
| 5043904 | 25 | 5043904CD1 | 59 | 5043904CB1 |
| 5202390 | 26 | 5202390CD1 | 60 | 5202390CB1 |
| 5526375 | 27 | 5526375CD1 | 61 | 5526375CB1 |
| 5677408 | 28 | 5677408CD1 | 62 | 5677408CB1 |
| 5982278 | 29 | 5982278CD1 | 63 | 5982278CB1 |
| 6437362 | 30 | 6437362CD1 | 64 | 6437362CB1 |
| 4173970 | 31 | 4173970CD1 | 65 | 4173970CB1 |
| 2772751 | 32 | 2772751CD1 | 66 | 2772751CB1 |
| 2793768 | 33 | 2793768CD1 | 67 | 2793768CB1 |
| 3035248 | 34 | 3035248CD1 | 68 | 3035248CB1 |

Table 2

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: | Probability score | GenBank Homolog |
|---------------------------|-----------------------------|-------------------|----------------------|---|
| 1 | 1889577CD1 | g3347848 | 0.00E+00 | kinesin light chain 2 [Mus musculus] |
| 2 | 2427982CD1 | g2760161 | 3.00E-64 | outer arm dynein light chain 2 [Anthocidaris crassispina] |
| 3 | 2470833CD1 | g11094032 | 1.00E-147 | [Mus musculus] (AF312712) gamma-parvin |
| 4 | 2080579CD1 | g11036542 | 0 | [Homo sapiens] (AF237772) gamma-parvin |
| | | g6141549 | 2.50E-101 | JNK/SAPK-associated protein-1 (JIP-1) |
| 5 | 2156553CD1 | g5419859 | 2.00E-170 | scaffold protein [Mus musculus] (Meyer, D. et al. (1999) J. Biol. Chem. 574:35113-35118) |
| 6 | 2182855CD1 | g2276319 | 0 | hypothetical protein similar to tubulin- |
| 7 | 2242106CD1 | g3834443 | 2.00E-13 | tyrosine ligase [Homo sapiens] (Lafanechere, L. et al. (1998) J. Cell Sci. 11:171-181) |
| | | g18156 | 1.20E-10 | axonemal dynein heavy chain [Homo sapiens] |
| 8 | 2726877CD1 | g4778 | 1.30E-12 | [Drosophila melanogaster] cytoplasmic dynein intermediate chain isoform DIC5b |
| 9 | 2738233CD1 | g4185884 | 7.70E-33 | 70kD dynein intermediate chain [Chlamydomonas reinhardtii] |
| 10 | 1833116CD1 | g10880797 | 0 | Usol protein [Saccharomyces cerevisiae] (Nakajima, H. et al. (1991) J. Cell Biol. 113:245-260) |
| 11 | 1799CD1 | g3283070 | 1.70E-07 | Groovin (Kakapo) [Drosophila melanogaster] (Strumpf, D. and T. Volk (1998) J. Cell Biol. 143:1259-1270) |
| | | g12082089 | 0 | [Mus musculus] Syne-1a |
| | | g12082091 | 0 | [Homo sapiens] hARPX |
| | | g3283070 | 1.70E-07 | [Gallus gallus] GARPX |
| 12 | 119814CD1 | g3005599 | 5.00E-09 | p80 katanin [Xenopus laevis] (McNally, F.J., Thomas, S. (1998) Mol. Biol. Cell 9:1847- 1861) |
| 13 | 1295420CD1 | g3243131 | 4.40E-18 | [Homo sapiens] (AF052432) katanin p80 subunit |
| | | g5870837 | 1.00E-113 | titin [Drosophila melanogaster] (Machado, C. et al. (1998) J. Cell Biol. 141:321-333) |
| | | g180622 | 5.60E-37 | [Homo sapiens] titin-like protein |
| 14 | 1309364CD1 | g12667401 | 0 | cytoplasmic linker protein-170 alpha-2 [Homo sapiens] (Pierre, P. et al. (1992) Cell 70:887-900) |
| | | g12667403 | 0 | [Homo sapiens] NUF2R |
| | | | 0 | [Mus musculus] NUF2R |

Table 2 (cont.)

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: | Probability score | GenBank Homolog |
|---------------------------|-----------------------------|-------------------|----------------------|--|
| 15 | 1315267CD1 | g53996 | 8.00E-74 | Tcp-10 (transmission control protein) [Mus musculus] (Davies, P. et al. (1991) Mamm. Genome 1:235-241) |
| 16 | 1403289CD1 | g5733814 | 4.60E-196 | angiotensin II AT2 receptor-interacting protein (Bedecs, K. et al. (1997) Biochem. J. 325:449-454) |
| 17 | 1607607CD1 | g3158498 | 1.60E-19 | Contains similarity to Pfam domain: PF00628 (PHD finger) (Aasland, R. et al. (1995) Trends Biochem. Sci. 20:56-59) |
| 18 | 1660025CD1 | g3253105 | 9.80E-20 | [Caenorhabditis elegans] strong similarity to the SNF2/RAD54 family of helicases (Eisen, J. et al. (1995) Nucleic Acids Res. 23:2715-2723) |
| 19 | 1796836CD1 | g414111 | 7.20E-14 | class II INCENP protein (inner centromere protein) [Gallus gallus] (Mackay, A. et al. (1993) J. Cell Biol. 123:373-385) |
| 20 | 2880670CD1 | g1813638 | 6.90E-16 | PF20 [Chlamydomonas reinhardtii] (Smith, E. and P. Lefebvre (1997) Mol. Biol. Cell 8:455-467) |
| 21 | 2913976CD1 | g63898 | 3.10E-56 | Zyxin [Gallus gallus] (Sadler, I. et al. (1992) J. Cell Biol. 119:1573-1587) |
| 22 | 3092084CD1 | g1154645 | 2.30E-10 | head-elevated expression in 0.9 kb [Drosophila melanogaster] (Yang, M.Y. et al. (2000) Genetics 154:285-297) |
| 23 | 3882482CD1 | g5825592 | 7.60E-171 | katanin p60 [Xenopus laevis] |
| 24 | 4933451CD1 | g684936 | 3.20E-30 | peptide with resemblance to the actin family [Homo sapiens] |
| 25 | 5043904CD1 | g2832237 | 2.10E-06 | cep250 centrosome associated protein [Homo sapiens] (Mack, G.J. et al. (1998) Arthritis Rheum. 41:551-558) |
| 26 | 5202390CD1 | g6572155 | 2.90E-21 | [Homo sapiens] dJ1014D13.2 (novel protein similar to ACTN3 (actinin, alpha 3)) |
| 27 | 5526375CD1 | g2443272 | 2.80E-77 | motor domain of KIF12 [Mus musculus] (Nakagawa, T. et al. (1997) Proc. Natl. Acad. Sci. USA 94:9654-9659) |
| 28 | 5677408CD1 | g6651427 | 2.20E-05 | dynein light intermediate chain 1 (LIC-2) [Rattus norvegicus] (Hughes, S.M. et al. (1995) J. Cell Sci. 108:17-24) |
| 29 | 5982278CD1 | g6006743 | 0 | mitotic kinesin-like protein 1 [Danio rerio] |
| | | g6723675 | 0 | [Homo sapiens] mitotic kinase-like protein-1 |

Table 2 (cont.)

| Polypeptide SEQ ID NO: | Incyte Polypeptide ID | GenBank ID NO: | Probability score | GenBank Homolog |
|---------------------------|-----------------------------|-------------------|----------------------|--|
| 30 | 6437362CD1 | g4929268 | 2.00E-38 | LOMP protein (LIM and PDZ domain protein) [Homo sapiens] |
| 31 | 4173970CD1 | g3879121 | 1.10E-146 | predicted using Genefinder; Similarity to Mouse ankyrin [Caenorhabditis elegans] |
| 32 | 2772751CD1 | g4545313 | 7.00E-103 | [Mus musculus] prominin-like protein (Corbeil, D. et al. (2000) J. Biol. Chem. 275:5512-5520) |
| 33 | 2793768CD1 | g485107 | 2.70E-85 | weakly similar to ANK repeat region of Fowlpox virus BamHI-orf7 |
| 34 | 3035248CD1 | g1333846 | 8.20E-45 | intermediate filament associated protein [Cricetulus griseus] (Skalli, O. et al. (1994) J. Cell Biol. 125:159-170) |

Table 3

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|------------|-----------------------|---------------------|---|-------------------------------|---|---|
| 1 | 1889577CD1 | 622 | T30 S90 T451 S499 S507 S539 T568 S615 Y345 Y431 S428 S557 S581 S619 S13 S151 T163 S232 T470 S507 S519 S521 T568 S589 S610 | N449 N587 | Kinesin light chain repeat BL01160:V88-S141, G191-P237, D238-A266, L267-C305, A308-R348, R349-C375, Q379-E420, E433-K480, L12-P50 KINESIN LIGHT CHAIN SIGNATURE PR00381:A97-A114, G191-S210, R213-T231, H278-R295, D322-E342, R357-K378 KINESIN LIGHT CHAIN PROTEIN KLC MOTOR MICROTUBULES COILED COIL REPEAT PD012762:L12-Q174 KINESIN LIGHT CHAIN REPEAT DM01439 A41539 1-234:M1-L220 signal_cleavage:M1-P50 Kinesin light chain repeat Kinesin2:Q223-N264, D265-K306 Kinesin light chain repeat Kinesin_Light:Q223-N264 D265-K306 Leucine Rich Repeat LRR:N49-K70, N 71-G92, T94-K115, K116-P140 COSMID C06A8 PROTEIN T09A5.9 CHROMOSOME III LEUCINEREPEAT REPEAT PD035408:S54-T179 Leucine Rich Repeat signature PR00019:L69-I82 | BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO SPScan HMMER-PFAM MOTIFS HMMER-PFAM BLAST-PRODOM BLIMPS-PRINTS |
| 2 | 2427982CD1 | 190 | T6 T94 T6 T167 | | | |
| 3 | 2470833CD1 | 331 | S37 S67 T188 S267 S293 T36 S37 S101 S169 S176 T188 T305 T317 | N55 N114 N274 | Calponin homology (CH) domain CH:N210-T317 | HMMER-PFAM |
| 4 | 2080579CD1 | 239 | T92 S148 T174 S191 S26 | | | MOTIFS |
| 5 | 2156553CD1 | 488 | S237 T370 T402 T121 T226 S428 | N167 N168 | PROTEIN CHROMOSOME TUBULIN TYROSINE LIGASE TTL C55A6.2 ZK1128.6 III PD008766:G63-V285 | BLAST-PRODOM |

Table 3 (cont.)

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|------------|-----------------------|---------------------|---|---------------------------------------|---|--|
| 6 | 2182855CD1 | 1190 | S26 S197 S207 S312 T354 S356 S492 S509 S519 S593 S686 T902 S920 S976 T992 Y811 Y955 T45 T98 S149 T163 T233 T350 S406 T446 S468 S524 S896 S976 | N20 N23 N156 N308 N433 N548 N635 N777 | PROTEIN DYNEIN CHAIN MOTOR MICROTUBULES ATPBINDING HEPTAD REPEAT PATTERN HEAVY PD003982:S920-V1190 do DYNEIN; HEAVY; CILIARY; CYTOSOLIC DM04585 P39057 2948-4465:L14-V1190 G_Beta_Repeats:L130-N144 | BLAST-PRODOM BLAST-DOMO MOTIFS |
| 7 | 2242106CD1 | 270 | S15 S56 S168 S189 S203 T240 S77 T213 | | signal_cleavage:M1-T25 transmem_domain:V29-L53 WD domain, G-beta repeat WD40:A116-S155, T207-Q245 | SPScan HMMER HMMER-PFAM |
| 8 | 2726877CD1 | 647 | S38 T173 T184 S322 S442 T483 T503 S510 S589 T14 S122 T134 T189 T408 S447 S461 S472 S510 S579 S593 | N99 N120 N316 N480 N508 N644 | PROTEIN COILED COIL CHAIN MYOSIN REPEAT HEAVY ATPBINDING FILAMENT HEPTAD PD000002:Q409-E619 | BLAST-PRODOM |
| 9 | 2738233CD1 | 1086 | T386 S12 S32 T86 S142 T251 S298 T343 T404 T414 T421 S427 S512 S559 S594 S618 T651 S671 T748 T799 S825 S870 S900 S954 S962 S963 Y146 S58 S75 S185 T286 S307 S366 T404 S512 S556 S559 S658 S675 S977 T987 | N182 N359 N545 | Spectrin repeat:R2-E66, N69-E171, V174-E285, R288-H394, G397-R501, T699-Q726, Q729-D836 Spectrin repeat proteins PF00435:W155-K170 | HMMER-PFAM BLIMPS-PFAM |
| 10 | 1833116CD1 | 396 | S104 S108 T341 T343 S367 T378 S388 T34 S163 S189 T243 S258 | N21 N101 | P59-K245, Q266-D395 Actin T3-L37, Q73-A127, R137-R191, I289-T343, D346-D395 Actins V333-K393 Actins signatures L4-I289 ACTINS AND ACTIN-RELATED PROTEINS Q266-K393 ACTIN | HMMER_PFAM BLIMPS_BLOCKS PROFILES SCAN BLAST_DOMO BLAST_PRODOM |

Table 3 (cont.)

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|------------|-----------------------|---------------------|--|-------------------------------|--|---|
| 11 | 001799CD1 | 304 | Y121 Y183 T33 S43 T58 S137 S254 S30 S89 S176 S131 S229 S255 | N31 N52 N124 | W172-D189 Aldehyde dehydrogenases motif V147-W249 P_value 5.9e-07 KATANIN P80 centrosome-binding subunit | BLIMPS_BLOCKS BLAST_PRODOM |
| 12 | 119814CD1 | 201 | T139 Y24 S4 S60 S68 T96 T106 T144 | N148 | M1-A23 signal cleavage G66-T125, S10-A28 Immunoglobulin domain | SPSCAN HMMER_PFAM |
| 13 | 1295420CD1 | 547 | T399T2 T58 T77 T192 T260 S491 T518 T154 T309 T374 T377 T386 T454 T515 T518 | N190 | G314-F345 G436-F467 Cap_Gly G314-S356, G436-P478 CAP-Gly domain T117-R158, T160-S191, N197-R229 Ank repeat G321-F345 CAP-Gly domain proteins G436-F476 MICROTUBULES CYTOSKELETON COILED COIL | MOTIFS HMMER_PFAM HMMER_PFAM BLIMPS_BLOCKS BLAST_PRODOM |
| 14 | 1309364CD1 | 464 | Y369 Y445 S118 T117 S217 T220 S232 S239 S340 T24 T32 T90 T137 S147 S232 T372 T428 | N30 N215 | E417-P492, L294-K362 CAP(cytoskeleton-associated protein)-GLY DOMAIN Q177-K418 COILED COIL MYOSIN REPEAT | BLAST_DOMO BLAST_PRODOM |
| 15 | 1315267CD1 | 569 | S3 T68 S85 S103 T229 S306 S356 T408 T482 T535 S551 T246 S20 T29 T31 T146 S167 T217 S292 T318 S385 T450 T477 | N46 N121 N155 N304 N406 | I465-E475 Muscarinic M4 receptor signature H386-A524 TCOMPLEX Male germ-cell specific protein | BLIMPS_PFAM BLAST_PRODOM |
| 16 | 1403289CD1 | 436 | S4 T17 S111 T167 T212 S222 S255 S308 S407 S421 S434 T35 S390 T58 T76 T97 T139 T153 T187 S213 S220 S235 S249 S270 Y74 | N80 N336 | L254-L275 L306-L327 Leucine zipper K196-E395 COILED COIL MYOSIN REPEAT Q107-E395 TRICHOHYALIN (hair root sheath protein) | MOTIFS BLAST_PRODOM BLAST_DOMO |

Table 3 (cont.)

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|------------|-----------------------|---------------------|--|-------------------------------|--|--|
| 17 | 1607607CD1 | 363 | S106 T206 T275 S288 S324 T331 T341 S51 T63 T162 S212 T336 | N229 N307 | L6-L27 L55-L76 Leucine Zipper L6-El6 Prepro orexin signature S3-El99 F33E11.3 PROTEIN similar to PHD finger K159-K165 Regulator of G protein signaling domain Q92-R103 5-hydroxytryptamine 2C receptor signature E60-P247 TOPOISOMERASE I DNA ISOMERASE REPEAT A17-S246 CYLICIN II sperm head cytoskeletal protein Q133-K383 COILED COIL MYOSIN REPEAT Q135-Q412 TRICHOHYALIN (hair root sheath protein) | MOTIFS BLIMPS_PRINTS BLAST_PRODUM BLIMPS_PPFAM BLIMPS_PRINTS BLAST_PRODUM BLAST_DOMO BLAST_PRODUM BLAST_DOMO |
| 18 | 1660025CD1 | 247 | S366 S45 S69 T96 S139 S148 S161 S183 S238 T264 T392 S416 Y399 S224 T264 T369 S381 | | WDREPEAT PROTEIN PF20 REPEAT WD FLAGELLA PD134845: E51-A178 LIM domains: C22-E80; C82-A139; C142-A208 LIM domain BL00478: Y43-L57 LIM domain signatures: E3-Y75; Y63-R206; M1-K137 LIM METAL-BINDING REPEAT DM00055 Q04584 464-533: F134-H203 LIM domain motifs: C22-L57; C82-I115; C142-L181 CALDESMON DM06224 P12957 1-755: T7-N197 (P-value = 7.6e-08) | BLAST_PRODUM HMMER-PFAM BLIMPS-BLOCKS PROFILES SCAN BLAST-DOMO MOTIFS BLAST-DOMO |
| 19 | 1796836CD1 | 441 | S366 S45 S69 T96 S139 S148 S161 S183 S238 T264 T392 S416 Y399 S224 T369 S381 | N113 N128 | | |
| 20 | 2880670CD1 | 183 | T48 S53 S68 T88 | N66 | | BLAST_PRODUM |
| 21 | 2913976CD1 | 212 | S124 S143 T49 T52 Y75 Y172 | | | HMMER-PFAM BLIMPS-BLOCKS PROFILES SCAN |
| 22 | 3092084CD1 | 227 | T11 S79 S56 S58 S113 S221 S6 T7 S46 T201 Y141 | N19 | | BLAST-DOMO MOTIFS BLAST-DOMO |

Table 3 (cont.)

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|------------|-----------------------|---------------------|---|-------------------------------|---|---|
| 23 | 3882482CD1 | 490 | T318 S42 S71 S87 S143 T150 S174 T318 S416 S440 S448 T457 S473 S75 S145 T252 T278 S285 S314 S325 T382 Y377 | N271 | AAA ATP-binding protease domain: G243-R433 AAA-protein family protease signatures BL00674: W207-P227; W241-A262; S274-R316 G338-K384; G414-R433 KATANIN P60 SUBUNIT PD116869: M1-P135 AAA-PROTEIN FAMILY DM00024 P34808 188-348: D208-L368 AAA motif: V352-R370 ATP/GTP binding site (P-loop): G248-T255 Actin domain: M1-M114 Actins proteins signature BL00406: T5-K39 Actin signature PR00190: E24-V33 PROTEIN STRUCTURAL ACTIN MULTIGENE FAMILY ACETYLATION MUSCLE CYTOSKELETON CYTOPLASMIC ACTINLIKE PD000056: V6-L117 ACTINS AND ACTIN-RELATED PROTEINS DM00167 P20360 3-272: A2-M114 PROTEIN COILED COIL CHAIN MYOSIN REPEAT HEAVY ATPBINDING FILAMENT HEPTAD PD000002: L617-D679 (P-value = 2.5e-05) C2-DOMAIN DM00150 P24506 157-283: R460-S584 (P-value = 7.2e-06) | HMME-PFAM BLIMPS-BLOCKS BLAST-PRODOM BLAST-DOMO MOTIFS MOTIFS HMME-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO |
| 24 | 4933451CD1 | 133 | S52 S115 | | | |
| 25 | 5043904CD1 | 912 | T162 S592 S773 T28 S54 S63 S81 S135 S251 S260 S278 T299 T374 S383 S553 S565 T620 T636 S647 T690 S744 T786 S823 S880 S889 S908 S7 T50 S157 T192 S199 S243 T308 T326 S334 T542 T550 S576 S584 S640 T671 S752 S766 S774 S777 S780 T862 Y681 | N483 N742 | | BLAST-DOMO |

Table 3 (cont.)

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|------------|-----------------------|---------------------|--|-------------------------------------|--|---|
| 26 | 5202390CD1 | 1076 | S602 S903 T10 T45 T176 S204 T206 S393 S422 S454 S469 T475 T492 S494 S509 S518 S575 S592 T630 S632 S657 S732 S748 T759 T776 S822 S828 S880 T954 S982 S22 S57 S140 S184 T185 S203 T238 S253 S273 S316 S406 T410 S505 S548 T606 S626 S647 T835 T842 T856 T885 T898 T950 Y149 Y198 Y733 Y918 | N150 N289 N312 N405 N421 N462 | Calponin homology (CH) domain: P288-S393 ALPHA-ACTININ ACTIN-BINDING DOMAIN DM00325 P18091 28-252: A290-L385 | HMME-PFAM BLAST-DOMO |
| 27 | 5526375CD1 | 542 | S223 S313 S36 T80 S162 S280 S113 T158 S179 S218 S231 S272 T355 T366 S540 | N188 N292 | Kinesin motor domain: R31-N394 Kinesin motor domain BL00411: P25-E39; G100-G121; V157-L175 G216-L240; L264-L305; H314-P344 Kinesin motor domain signature: A247-A297 Kinesin heavy chain signatures PR00380: G100-G121; H225-I242; K263-R281; I315-T336 PROTEIN MOTOR ATPBINDING COILED COIL MICROTUBULES KINESINLIKE KINESIN MITOSIS HEAVY PD000458: R31-L401 KINESIN MOTOR DOMAIN DM00198 P46871 3-343: E23-Q370 ATP/GTP binding site (P-loop): G109-T116 Kinesin motor domain motif: G262-E273 ATP/GTP binding site (P-loop): G38-T45 | HMME-PFAM BLIMPS-BLOCKS PROFILESAN BLIMPS-PRINTS BLAST-PRODOR BLAST-DOMO MOTIFS MOTIFS MOTIFS |
| 28 | 5677408CD1 | 351 | S340 T5 S22 T77 S92 S136 S186 S221 S284 T304 T105 S196 T205 T220 S343 Y336 | N20 | | |

Table 3 (cont.)

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|------------|-----------------------|---------------------|--|-----------------------------------|---|---|
| 29 | 5982278CD1 | 856 | T13 S18 S807 S808 T84 S125 T200 S221 T266 S272 S334 T347 T369 T580 S710 T757 S763 T789 T823 S3 T8 S74 S116 S186 S187 S298 S341 T344 T366 T382 S605 T622 T663 T679 T793 S802 | N48 N49 N345 N361 N565 N800 | Kinesin motor domain: R31-E466 Kinesin motor domain BL00411: P25-D39; K69-Q85; G103-G124 G130-F140; Y216-L234; G283-I307 L333-L374; N385-P415 Kinesin motor domain signatures: D317-L364 Kinesin heavy chain signatures PR00380: G103-G124; T292-L309; Q332-E350; V386-V407 PROTEIN MOTOR MICROTUBULES ATPBINDING COILED COIL KINESIN LIKE SIMILAR MITOTIC PROTEIN1 PD013891: E664-L841 KINESIN MOTOR DOMAIN DM00198 Q02241 4-443: A4-I446 ATP/GTP binding site (P-loop): G112-T119 Kinesin motor domain motif: S331-E342 LIM domain: C986-S1049 Calponin family repeat BL01052: F44-I69 LIM domain: F1008-L1022 LIM domain signature: K964-S1047 Calponin signature PR00889: S25-F39; C55-L72 CALPONIN FAMILY REPEAT DM01491 P51911 6-147: P4-S116 LIM domain motif: C986-L1022 | HMMER-PFAM BLIMPS-BLOCKS PROFILES CAN BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO MOTIFS MOTIFS HMMER-PFAM BLIMPS-BLOCKS BLIMPS-BLOCKS PROFILES CAN BLIMPS-PRINTS BLAST-DOMO MOTIFS |
| 30 | 6437362CD1 | 1056 | T164 S256 S377 S408 S979 S38 T46 S173 S231 S233 S327 T411 S435 T451 S511 S601 S637 S746 S836 S846 S897 S951 S972 S1027 T35 S109 T113 S120 S169 S185 S251 S373 S403 T529 T537 T553 T630 S647 S702 S710 S753 S889 S974 S981 Y179 Y294 | N272 N275 N475 N609 | Ank repeat: K347-N379; N77-K109; K110-E142; T144-V176; S177-G209; L212-E244; N246-K278; T279-V311; S314-K346 signal_cleavage: M1-A18 signal_peptide: M1-A18 transmembrane domain: V2-N22, V277-G296, L326-V344 AC133 ANTIGEN PROMININ HOMOLOG: M1-R333 | HMMER-PFAM SPSCAN HMMER HMMER BLAST_PRODOM |
| 31 | 4173970CD1 | 1569 | | | | |
| 32 | 2772751CD1 | 680 | S113 S212 S340 S343 S357 T104 T159 T276 T328 T342 | N210 N233 N251 N56 N66 N89 | | |

Table 3 (cont.)

| SEQ ID NO: | Incyte Polypeptide ID | Amino Acid Residues | Potential Phosphorylation Sites | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs | Analytical Methods and Databases |
|------------|-----------------------|---------------------|--|-------------------------------|---|---|
| 33 | 2793768CD1 | 590 | S3 S373 S432 S461 S495 S499 S548 S56 S561 S586 S83 T199 T228 T234 T255 T286 T331 T354 T357 T433 T445 T454 T534 T547 T588 | N101 N166 N233 | signal_cleavage: M1-S56 Ankyrin repeat: R40-R72, Q73-T102 SIMILAR TO ANKYRIN REPEAT REGION OF FOWLPOX VIRUS BAMHIORF7 PROTEIN: W75-H284, G348-Y469 | SPSCAN HMMER_PFAM BLAST_PRODUM |
| 34 | 3035248CD1 | 315 | S151 S291 T115 T207 T273 | | signal_cleavage: M1-G26 signal_peptide: M1-S24 transmembrane domain: P4-F20 INTERMEDIATE FILAMENT ASSOCIATED PROTEIN K147-T218 TROPOMYOSIN ALTERNATIVE SPLICING SIGNAL PRECURSOR CHAIN: E39-S213 INTERMEDIATE FILAMENT: K121-T218 | SPSCAN HMMER HMMR BLAST_PRODUM BLAST_PRODUM BLAST_DOMO |

Table 4

| Polynucleotide SEQ ID NO: | Incyte Polynucleotide ID | Sequence Length | Selected Fragment(s) | Sequence Fragments | 5' Position | 3' Position |
|------------------------------|--------------------------------|--------------------|---|--------------------------|-------------|-------------|
| 35 | 1889577CB1 | 2345 | 902-952 | 3824958H1 (BRAINOT01) | 1 | 284 |
| 35 | 1889577CB1 | 2345 | | 1915360R6 (PROSTUT04) | 774 | 1278 |
| 35 | 1889577CB1 | 2345 | | 1812980F6 (PROSTUT12) | 102 | 646 |
| 35 | 1889577CB1 | 2345 | | 3152565H1 (ADREN004) | 1799 | 2066 |
| 35 | 1889577CB1 | 2345 | | 1369763H1 (BSTMN002) | 1608 | 1840 |
| 35 | 1889577CB1 | 2345 | | 1784544F6 (BRAINOT10) | 1873 | 2345 |
| 35 | 1889577CB1 | 2345 | | 674685H1 (CRBLNOT01) | 1516 | 1769 |
| 35 | 1889577CB1 | 2345 | | 2838122F6 (DRGLNOT01) | 1016 | 1563 |
| 35 | 1889577CB1 | 2345 | | 1649402F6 (PROSTUT09) | 384 | 896 |
| 36 | 2427982CB1 | 709 | | 71263527V1 | 624 | 709 |
| 36 | 2427982CB1 | 709 | | 71247061V1 | 1 | 683 |
| 37 | 2470833CB1 | 1569 | 1-721 | 1684180F6 (PROSNOT15) | 1 | 497 |
| 37 | 2470833CB1 | 1569 | | 868966R6 (LUNGAST01) | 842 | 1407 |
| 37 | 2470833CB1 | 1569 | | 3576193T6 (BRONNOT01) | 980 | 1569 |
| 37 | 2470833CB1 | 1569 | | 1534629F1 (SPLNNOT04) | 677 | 1197 |
| 37 | 2470833CB1 | 1569 | | 5296329H1 (COLENOT02) | 493 | 730 |
| 37 | 2470833CB1 | 1569 | | 1716065F6 (UCMCNOT02) | 207 | 652 |
| 38 | 2080579CB1 | 1172 | 1-148, 686-854 | 868135H1 (BRAITUT03) | 368 | 632 |
| 38 | 2080579CB1 | 1172 | | 3458305F6 (293TF1T01) | 1 | 433 |
| 38 | 2080579CB1 | 1172 | | 2080579T6 (UTRSNOT08) | 583 | 1143 |
| 38 | 2080579CB1 | 1172 | | 2361824H1 (LUNGFT05) | 939 | 1172 |
| 38 | 2080579CB1 | 1172 | | 5174845H1 (EPIBFT01) | 442 | 637 |
| 39 | 2156553CB1 | 2380 | 1-360, 2126-2380, 1121-1655 | 5322363H1 (FIBPFEN06) | 1 | 260 |
| 39 | 2156553CB1 | 2380 | | 2916949T6 (THYMFET03) | 1769 | 2380 |
| 39 | 2156553CB1 | 2380 | | 865038X304D1 (BRAITUT03) | 371 | 937 |
| 39 | 2156553CB1 | 2380 | | 2916949F6 (THYMFET03) | 1341 | 1886 |
| 39 | 2156553CB1 | 2380 | | 2156553F6 (BRAINOT09) | 705 | 1224 |
| 39 | 2156553CB1 | 2380 | | 1722673F6 (BLADNOT06) | 1176 | 1582 |
| 39 | 2156553CB1 | 2380 | | 4899945H1 (OVARIT01) | 245 | 531 |
| 39 | 2156553CB1 | 2380 | | 1758833H1 (PITUNOT03) | 236 | 485 |
| 40 | 2182855CB1 | 4396 | 2141-2667, 1- 1505, 1737- 1890, 3803- 4396 | 2816335F6 (BRSTNOT14) | 3897 | 4396 |
| 40 | 2182855CB1 | 4396 | | 2967273F6 (SCORNOT04) | 580 | 1116 |

Table 4 (cont.)

| Polynucleotide SEQ ID NO: | Incyte Polynucleotide ID | Sequence Length | Selected Fragment(s) | Sequence Fragments | 5' Position | 3' Position |
|------------------------------|--------------------------------|--------------------|--|---------------------------|-------------|-------------|
| 40 | 2182855CB1 | 4396 | | 1611084F6 (COLNTUT06) | 3554 | 3992 |
| 40 | 2182855CB1 | 4396 | | 1785722H1 (BRAINT010) | 2610 | 2857 |
| 40 | 2182855CB1 | 4396 | | 2182855F6 (SININT01) | 3389 | 3914 |
| 40 | 2182855CB1 | 4396 | | 1484284F6 (CORPNOT02) | 1 | 607 |
| 40 | 2182855CB1 | 4396 | | 92276318 | 61 | 3799 |
| 40 | 2182855CB1 | 4396 | | 2321435H1 (OVARNOT02) | 2609 | 2794 |
| 40 | 2182855CB1 | 4396 | | 1578313H1 (DUODNOT01) | 1452 | 1546 |
| 40 | 2182855CB1 | 4396 | | 1618459F6 (BRAITUT12) | 2925 | 3464 |
| 40 | 2182855CB1 | 4396 | | 2321435X308F1 (OVARNOT02) | 1811 | 2391 |
| 41 | 2242106CB1 | 1831 | 1-509, 626-1018 | 965728R1 (BRSTNOT05) | 1283 | 1831 |
| 41 | 2242106CB1 | 1831 | | 1650350F6 (PROSTUT09) | 322 | 1031 |
| 41 | 2242106CB1 | 1831 | | 1396324T1 (THYRNOT03) | 948 | 1660 |
| 41 | 2242106CB1 | 1831 | | 6843794H1 (KIDNTMN03) | 191 | 985 |
| 41 | 2242106CB1 | 1831 | | 956964T1 (KIDNNOT05) | 1019 | 1667 |
| 41 | 2242106CB1 | 1831 | | 70846228V1 | 1 | 216 |
| 42 | 2726877CB1 | 3249 | 1979-2045, 2854-2873, 1857-1919, 2543-2608 | 3728286F6 (SMCCNON03) | 1073 | 1504 |
| 42 | 2726877CB1 | 3249 | | 3645568F6 (LUNGNOT34) | 383 | 933 |
| 42 | 2726877CB1 | 3249 | | 4969912H1 (KIDEUNC10) | 151 | 425 |
| 42 | 2726877CB1 | 3249 | | 2500944T6 (ADRETUT05) | 1254 | 1839 |
| 42 | 2726877CB1 | 3249 | | 2726877F6 (OVARUT05) | 1603 | 2060 |
| 42 | 2726877CB1 | 3249 | | 4195125T6 (COLITUT02) | 2652 | 3249 |
| 42 | 2726877CB1 | 3249 | | 4972430H1 (HELATXT02) | 901 | 1191 |
| 42 | 2726877CB1 | 3249 | | 4195125F6 (COLITUT02) | 2257 | 2810 |
| 42 | 2726877CB1 | 3249 | | 5492146H1 (DRGTNON04) | 2128 | 2380 |
| 42 | 2726877CB1 | 3249 | | 3894803H1 (TLYMN05) | 1 | 296 |
| 42 | 2726877CB1 | 3249 | | 3728286T6 (SMCCNON03) | 1682 | 2279 |
| 43 | 2738233CB1 | 4133 | 1-194, 4026-4133, 1607-2971, 3066-3186 | 1649466F6 (PROSTUT09) | 2737 | 3394 |
| 43 | 2738233CB1 | 4133 | | 2267313R6 (UTRSNOT02) | 2211 | 2730 |
| 43 | 2738233CB1 | 4133 | | 93882312_CD | 126 | 3516 |
| 43 | 2738233CB1 | 4133 | | 2945874H2 (BRAITUT23) | 2710 | 3006 |
| 43 | 2738233CB1 | 4133 | | 2242201F6 (PANCITUT02) | 1538 | 2077 |

Table 4 (cont.)

| Polynucleotide SEQ ID NO: | Incyte Polynucleotide ID | Sequence Length | Selected Fragment(s) | Sequence Fragments | 5' Position | 3' Position |
|------------------------------|--------------------------------|--------------------|-------------------------|------------------------|-------------|-------------|
| 43 | 2738233CB1 | 4133 | | 653470R6 (EOSINOT03) | 685 | 1068 |
| 43 | 2738233CB1 | 4133 | | 2267313T6 (UTRSNOT02) | 3481 | 4109 |
| 43 | 2738233CB1 | 4133 | | 2962383H1 (ADRENOT09) | 1 | 277 |
| 43 | 2738233CB1 | 4133 | | 1467024F6 (PANCYTUT02) | 1948 | 2420 |
| 43 | 2738233CB1 | 4133 | | 3555108H1 (SYNONOT01) | 1394 | 1704 |
| 43 | 2738233CB1 | 4133 | | 2965786H1 (SCORNOT04) | 217 | 469 |
| 43 | 2738233CB1 | 4133 | | 1262951R1 (SYNORAT05) | 3959 | 4133 |
| 43 | 2738233CB1 | 4133 | | 1260043T1 (MENITUT03) | 3370 | 4017 |
| 43 | 2738233CB1 | 4133 | | 1486351H1 (CORENOT02) | 1236 | 1463 |
| 43 | 2738233CB1 | 4133 | | 3765643F6 (BRSTNOT24) | 277 | 758 |
| 44 | 1833116CB1 | 1754 | 1700-1754 | 2852676F6 (BRSTTUT13) | 25 | 529 |
| 44 | 1833116CB1 | 1754 | | 3128645H1 (LUNGNOT12) | 539 | 839 |
| 44 | 1833116CB1 | 1754 | | 5016828H1 (BRAXNOT03) | 632 | 886 |
| 44 | 1833116CB1 | 1754 | | 413418R1 (BRSTNOT01) | 979 | 1615 |
| 44 | 1833116CB1 | 1754 | | 1442616R1 (THYRNOT03) | 1247 | 1754 |
| 44 | 1833116CB1 | 1754 | | 1785591H1 (BRAINOT10) | 1 | 284 |
| 44 | 1833116CB1 | 1754 | | 5172858H1 (EPITXT01) | 331 | 588 |
| 44 | 1833116CB1 | 1754 | | 1920612R6 (BRSTTUT01) | 866 | 1402 |
| 45 | 001799CB1 | 2713 | 1-27, 1464-2008 | 5994129H1 (FTUBTUT02) | 587 | 912 |
| 45 | 001799CB1 | 2713 | | 4245126H1 (BRABDIT01) | 1982 | 2239 |
| 45 | 001799CB1 | 2713 | | 6818763J1 (BRAUNOR01) | 1 | 549 |
| 45 | 001799CB1 | 2713 | | 5054327H1 (COLATMT01) | 1456 | 1726 |
| 45 | 001799CB1 | 2713 | | 6739739H1 (BRAFDIT02) | 2153 | 2713 |
| 45 | 001799CB1 | 2713 | | 71336820V1 | 276 | 898 |
| 45 | 001799CB1 | 2713 | | 3730557H1 (SMCCNON03) | 884 | 1198 |
| 45 | 001799CB1 | 2713 | | 644891R6 (BRSTTUT02) | 1575 | 1981 |
| 45 | 001799CB1 | 2713 | | 3515211H1 (LUNGNOT33) | 1940 | 2217 |
| 45 | 001799CB1 | 2713 | | 2691467T6 (LUNGNOT23) | 907 | 1471 |
| 45 | 001799CB1 | 2713 | | 3240741H1 (COLAUCT01) | 1401 | 1685 |
| 45 | 001799CB1 | 2713 | | 4771110H1 (BRATNOT02) | 1715 | 1990 |
| 46 | 119814CB1 | 1768 | 1-688 | 2637776F6 (BONTNOT01) | 958 | 1494 |
| 46 | 119814CB1 | 1768 | | 119814R1 (MUSCNOT01) | 739 | 1406 |
| 46 | 119814CB1 | 1768 | | 2638913F6 (BONTNOT01) | 1410 | 1768 |
| 46 | 119814CB1 | 1768 | | 1993563H1 (CORENOT02) | 656 | 940 |
| 46 | 119814CB1 | 1768 | | 92184959 | 349 | 846 |
| 46 | 119814CB1 | 1768 | | 91218792 | 1 | 579 |
| 46 | 119814CB1 | 1768 | | 2395927T6 (THPLAZT01) | 20 | 182 |

Table 4 (cont.)

| Polynucleotide SEQ ID NO: | Incyte Polynucleotide ID | Sequence Length | Selected Fragment (s) | Sequence Fragments | 5' Position | 3' Position |
|------------------------------|--------------------------------|--------------------|----------------------------|-----------------------|-------------|-------------|
| 47 | 1295420CB1 | 3287 | 914-1321, 177-207 | 148434H1 (FIBRNGT01) | 812 | 1028 |
| 47 | 1295420CB1 | 3287 | | 2697808H1 (UTRSNOT12) | 675 | 877 |
| 47 | 1295420CB1 | 3287 | | 5778293H1 (BRAXNOT03) | 1185 | 1447 |
| 47 | 1295420CB1 | 3287 | | 2149151F6 (BRAINOT09) | 1661 | 2176 |
| 47 | 1295420CB1 | 3287 | | 2883729F6 (SINJNOT02) | 1 | 515 |
| 47 | 1295420CB1 | 3287 | | 1888639H1 (BLADTUT07) | 868 | 1147 |
| 47 | 1295420CB1 | 3287 | | 2154634F6 (BRAINOT09) | 2720 | 3278 |
| 47 | 1295420CB1 | 3287 | | 3703932T6 (PENCNOT07) | 1334 | 2009 |
| 47 | 1295420CB1 | 3287 | | 998245R1 (KIDNTUT01) | 2118 | 2637 |
| 47 | 1295420CB1 | 3287 | | 2223877F6 (SEMVNOT01) | 323 | 816 |
| 47 | 1295420CB1 | 3287 | | 4027885H1 (BRAINOT23) | 999 | 1261 |
| 47 | 1295420CB1 | 3287 | | 835343R1 (PROSNOT07) | 2247 | 2814 |
| 47 | 1295420CB1 | 3287 | | 1367731R1 (SCORNON02) | 2856 | 3287 |
| 48 | 1309364CB1 | 1748 | 1-49, 1037-1135 | 4701455H1 (SMCRTXT01) | 1162 | 1416 |
| 48 | 1309364CB1 | 1748 | | 4904644F6 (TLYMNOT08) | 799 | 1399 |
| 48 | 1309364CB1 | 1748 | | 2914466F6 (THYMFET03) | 1 | 536 |
| 48 | 1309364CB1 | 1748 | | 5590953H1 (ENDINOT02) | 678 | 933 |
| 48 | 1309364CB1 | 1748 | | 1309364F6 (COLNFET02) | 1227 | 1748 |
| 48 | 1309364CB1 | 1748 | | 3727909H1 (SMCCNON03) | 513 | 816 |
| 49 | 1315267CB1 | 2163 | 705-799 | 898915H1 (BRSTTUT03) | 1839 | 2163 |
| 49 | 1315267CB1 | 2163 | | 465550R6 (LATRNOT01) | 673 | 1257 |
| 49 | 1315267CB1 | 2163 | | 1575785F6 (LNODNOT03) | 1029 | 1630 |
| 49 | 1315267CB1 | 2163 | | 5191222F6 (OVARDT06) | 1 | 538 |
| 49 | 1315267CB1 | 2163 | | 5207783F6 (BRAFNOT02) | 426 | 1041 |
| 49 | 1315267CB1 | 2163 | | 1315267F6 (BLADTUT02) | 1434 | 2022 |
| 50 | 1403289CB1 | 1615 | 1119-1170 | 2811792T6 (OVARNOT10) | 489 | 1121 |
| 50 | 1403289CB1 | 1615 | | 059048R6 (MUSCNOT01) | 1080 | 1615 |
| 50 | 1403289CB1 | 1615 | | 3502723H1 (ADRENOT11) | 932 | 1225 |
| 50 | 1403289CB1 | 1615 | | 1403289F6 (LATRTUT02) | 1 | 601 |
| 51 | 1607607CB1 | 1356 | 1-157, 1263-1356 | 7262994H1 (UTRETM01) | 252 | 929 |
| 51 | 1607607CB1 | 1356 | | 3467640F6 (BRAIDIT01) | 350 | 1048 |
| 51 | 1607607CB1 | 1356 | | 2137437H1 (ENDCNOT01) | 1 | 275 |
| 51 | 1607607CB1 | 1356 | | 1607607F6 (LUNGNOT15) | 872 | 1356 |
| 52 | 1660025CB1 | 1268 | 1-88, 424- 459, 775-836 | 2580277F6 (KIDNTUT13) | 213 | 932 |

Table 4 (cont.)

| Polynucleotide SEQ ID NO: | Incyte Polynucleotide ID | Sequence Length | Selected Fragment(s) | Sequence Fragments | 5' Position | 3' Position |
|------------------------------|--------------------------------|--------------------|-------------------------|--------------------------|-------------|-------------|
| 52 | 1660025CB1 | 1268 | | 2172241H1 (ENDNOT03) | 1 | 240 |
| 52 | 1660025CB1 | 1268 | | 1988094R6 (LUNGAST01) | 750 | 1268 |
| 52 | 1660025CB1 | 1268 | | 1756804R6 (PTUNOT03) | 248 | 966 |
| 53 | 1796836CB1 | 2554 | 1177-1245, 1-61 | 2280307T6 (PROSNON01) | 687 | 1302 |
| 53 | 1796836CB1 | 2554 | | 4971676H1 (HELATXT02) | 1 | 227 |
| 53 | 1796836CB1 | 2554 | | 2582430T6 (KIDNTUT13) | 1198 | 1545 |
| 53 | 1796836CB1 | 2554 | | 2497103T6 (ADRETUT05) | 1949 | 2533 |
| 53 | 1796836CB1 | 2554 | | 2534742H1 (BRAINOT18) | 1527 | 1758 |
| 53 | 1796836CB1 | 2554 | | 2553754T6 (THYNNOT03) | 1817 | 2528 |
| 53 | 1796836CB1 | 2554 | | 2726708H1 (OVARUT05) | 1630 | 1863 |
| 53 | 1796836CB1 | 2554 | | 276683H1 (TESTNOT03) | 1311 | 1568 |
| 53 | 1796836CB1 | 2554 | | 2938533H1 (THYMFET02) | 1056 | 1326 |
| 53 | 1796836CB1 | 2554 | | 6914750J1 (PTUDIR01) | 47 | 684 |
| 53 | 1796836CB1 | 2554 | | 2300549R6 (BRSTNOT05) | 2162 | 2554 |
| 53 | 1796836CB1 | 2554 | | 3030841F6 (HEARFET02) | 319 | 877 |
| 54 | 2880670CB1 | 1216 | 605-636 | 2889280T7 (LUNGFE04) | 489 | 1192 |
| 54 | 2880670CB1 | 1216 | | 1358092F1 (LUNGNOT09) | 246 | 920 |
| 54 | 2880670CB1 | 1216 | | 816703R1 (OVARUT01) | 657 | 1216 |
| 54 | 2880670CB1 | 1216 | | 2529604H1 (GBLANOT02) | 1 | 357 |
| 55 | 2913976CB1 | 1457 | 1-446, 1406-1457 | 3736188F6 (SMCCNOS01) | 1173 | 1428 |
| 55 | 2913976CB1 | 1457 | | 2913976F6 (KIDNTUT15) | 1 | 520 |
| 55 | 2913976CB1 | 1457 | | 4645636H1 (PROSTUT20) | 1187 | 1449 |
| 55 | 2913976CB1 | 1457 | | 4331439H1 (KIDNNOT32) | 461 | 712 |
| 55 | 2913976CB1 | 1457 | | 4643722H1 (PROSTMT03) | 1108 | 1325 |
| 55 | 2913976CB1 | 1457 | | 1312116F1 (COLNFET02) | 583 | 1127 |
| 56 | 3092084CB1 | 1636 | 857-1636 | 1709866F6 (PROSNOT16) | 1097 | 1634 |
| 56 | 3092084CB1 | 1636 | | 2807436F6 (BLADTUT08) | 1277 | 1636 |
| 56 | 3092084CB1 | 1636 | | 6906626H1 (MUSLTDR02) | 1 | 610 |
| 56 | 3092084CB1 | 1636 | | SEMA03169F1 | 597 | 1181 |
| 56 | 3092084CB1 | 1636 | | 3092084F6 (BRSTNOT19) | 543 | 1147 |
| 57 | 3882482CB1 | 1742 | 1-82, 923-994 | 2286328X19F1 (BRAINON01) | 742 | 1251 |
| 57 | 3882482CB1 | 1742 | | 2804312F6 (PENCNOT01) | 299 | 926 |
| 57 | 3882482CB1 | 1742 | | 986917T6 (LVENNOT03) | 1088 | 1742 |
| 57 | 3882482CB1 | 1742 | | 3175528F6 (UTRSTUT04) | 1 | 317 |
| 57 | 3882482CB1 | 1742 | | 1232503F6 (LUNGFE03) | 960 | 1553 |

Table 4 (cont.)

| Polynucleotide SEQ ID NO: | Incyte Polynucleotide ID | Sequence Length | Selected Fragment (s) | Sequence Fragments | 5' Position | 3' Position |
|------------------------------|--------------------------------|--------------------|--|--------------------------|-------------|-------------|
| 57 | 3882482CB1 | 1742 | | 2286328X16F1 (BRAINO01) | 402 | 969 |
| 58 | 4933451CB1 | 602 | 1-401 | 4931591H1 (BRSTTUT20) | 341 | 602 |
| 58 | 4933451CB1 | 602 | | 2502933F6 (CONUTUT01) | 1 | 479 |
| 59 | 5043904CB1 | 3237 | 1964-2007, 2911-2969, 993-1057, 1328-1589 | 2153280F6 (BRAINO09) | 335 | 888 |
| 59 | 5043904CB1 | 3237 | | 5043904R6 (PLACFER01) | 1 | 583 |
| 59 | 5043904CB1 | 3237 | | 075538H1 (THP1PEB01) | 1604 | 1818 |
| 59 | 5043904CB1 | 3237 | | 3189755X301D1 (THYMN004) | 1131 | 1674 |
| 59 | 5043904CB1 | 3237 | | 2460935F6 (THYRN008) | 614 | 1177 |
| 59 | 5043904CB1 | 3237 | | 1795345R6 (PROSTUT05) | 2026 | 2510 |
| 59 | 5043904CB1 | 3237 | | 1865353F6 (PROSN019) | 2686 | 3237 |
| 59 | 5043904CB1 | 3237 | | 1365052R6 (SCORN002) | 2440 | 2957 |
| 59 | 5043904CB1 | 3237 | | 3250182H1 (SEMVN003) | 1831 | 2128 |
| 59 | 5043904CB1 | 3237 | | 4713986H1 (BRAIHC01) | 2266 | 2513 |
| 59 | 5043904CB1 | 3237 | | 3804331H1 (BLADTUT03) | 1777 | 2089 |
| 60 | 5202390CB1 | 3640 | 1412-1489, 2294-2508, 1-161, 3244-3640 | 2544502H2 (UTRSN011) | 2656 | 2927 |
| 60 | 5202390CB1 | 3640 | | 2321656R6 (OVARN002) | 2039 | 2448 |
| 60 | 5202390CB1 | 3640 | | 2557486F6 (THYMN003) | 1 | 659 |
| 60 | 5202390CB1 | 3640 | | 1441193F6 (THYRN003) | 3317 | 3640 |
| 60 | 5202390CB1 | 3640 | | 2844888H1 (DRGLN001) | 864 | 1009 |
| 60 | 5202390CB1 | 3640 | | 3705507H1 (PENCN007) | 3047 | 3352 |
| 60 | 5202390CB1 | 3640 | | 2122377F6 (BRSTN007) | 2844 | 3297 |
| 60 | 5202390CB1 | 3640 | | 4313841F6 (BRAFN001) | 2322 | 2662 |
| 60 | 5202390CB1 | 3640 | | 5092148H1 (UTRSTMR01) | 685 | 961 |
| 60 | 5202390CB1 | 3640 | | 94240294 CD | 525 | 3562 |
| 60 | 5202390CB1 | 3640 | | 1944075H1 (PITUN001) | 1590 | 1851 |
| 60 | 5202390CB1 | 3640 | | 3493516H1 (ADRETUT07) | 1999 | 2279 |
| 60 | 5202390CB1 | 3640 | | 176099H1 (TYLNM001) | 411 | 738 |
| 60 | 5202390CB1 | 3640 | | 2851783H1 (BRSTTUT13) | 2600 | 2796 |
| 60 | 5202390CB1 | 3640 | | 534176R1 (BRAINO03) | 1041 | 1567 |
| 60 | 5202390CB1 | 3640 | | 4107451H1 (BRSTTUT17) | 1431 | 1694 |
| 60 | 5202390CB1 | 3640 | | 2842892F6 (DRGLN001) | 1604 | 2220 |

Table 4 (cont.)

| Polynucleotide SEQ ID NO: | Incyte Polynucleotide ID | Sequence Length | Selected Fragment(s) | Sequence Fragments | 5' Position | 3' Position |
|------------------------------|--------------------------------|--------------------|----------------------------------|---------------------------|-------------|-------------|
| 61 | 5526375CB1 | 2111 | 1-50, 1540-2111 | 3038423H1 (BRSTNOT16) | 1004 | 1278 |
| 61 | 5526375CB1 | 2111 | | 2513433F6 (LIVRTUT04) | 1600 | 2111 |
| 61 | 5526375CB1 | 2111 | | 3534575H1 (KIDNNOT25) | 1263 | 1529 |
| 61 | 5526375CB1 | 2111 | | 2993194F6 (KIDNFE02) | 729 | 1224 |
| 61 | 5526375CB1 | 2111 | | 5109694H1 (PROSTUS19) | 1385 | 1626 |
| 61 | 5526375CB1 | 2111 | | 2993263H1 (KIDNFE02) | 1242 | 1479 |
| 61 | 5526375CB1 | 2111 | | 3534157H1 (KIDNNOT25) | 518 | 788 |
| 61 | 5526375CB1 | 2111 | | 2580307F6 (KIDNTUT13) | 1 | 519 |
| 61 | 5526375CB1 | 2111 | | 2070882F6 (ISLTNOT01) | 225 | 715 |
| 62 | 5677408CB1 | 1389 | 1-177 | 535789R1 (ADREN0T03) | 1079 | 1389 |
| 62 | 5677408CB1 | 1389 | | 881149T6 (THYRN0T02) | 793 | 1378 |
| 62 | 5677408CB1 | 1389 | | 6023544H1 (TESTNOT11) | 736 | 1017 |
| 62 | 5677408CB1 | 1389 | | 881149R6 (THYRN0T02) | 1 | 772 |
| 63 | 5982278CB1 | 3331 | 809-1149, 1755-1989 | 5260541H1 (CONDUT01) | 2203 | 2470 |
| 63 | 5982278CB1 | 3331 | | 1390622H1 (EOSINOT01) | 2029 | 2257 |
| 63 | 5982278CB1 | 3331 | | 4515063H1 (EPIMNOT01) | 1740 | 1993 |
| 63 | 5982278CB1 | 3331 | | 3584113H1 (293TF4T01) | 2462 | 2796 |
| 63 | 5982278CB1 | 3331 | | 3405843H1 (ESOGNOT03) | 919 | 1174 |
| 63 | 5982278CB1 | 3331 | | 5261482H1 (CONDUT01) | 2611 | 2862 |
| 63 | 5982278CB1 | 3331 | | 1637273H1 (UTRSNOT06) | 1619 | 1733 |
| 63 | 5982278CB1 | 3331 | | 91521431 | 2559 | 3331 |
| 63 | 5982278CB1 | 3331 | | 3591491H1 (293TF5T01) | 1674 | 1981 |
| 63 | 5982278CB1 | 3331 | | 4957640H1 (TLYMNOT05) | 1464 | 1726 |
| 63 | 5982278CB1 | 3331 | | 4666088H1 (MEGBUNT01) | 2280 | 2536 |
| 63 | 5982278CB1 | 3331 | | 043258H1 (TLYN0T01) | 1370 | 1562 |
| 63 | 5982278CB1 | 3331 | | 2907496F6 (THYMN0T05) | 24 | 624 |
| 63 | 5982278CB1 | 3331 | | 4983673H1 (HELATXT05) | 1865 | 2134 |
| 63 | 5982278CB1 | 3331 | | 934671 CD | 130 | 3166 |
| 63 | 5982278CB1 | 3331 | | 3449505X304D1 (UTRSNON03) | 942 | 1450 |
| 63 | 5982278CB1 | 3331 | | 2640427T6 (LUNGUT08) | 2646 | 3309 |
| 63 | 5982278CB1 | 3331 | | 2205131F6 (SPINFE02) | 1 | 514 |
| 64 | 6437362CB1 | 3558 | 1-428, 3352-3558, 923-1583 | 720069R6 (SYNOOAT01) | 3159 | 3558 |
| 64 | 6437362CB1 | 3558 | | 2785980H1 (BRSTNOT13) | 2462 | 2730 |
| 64 | 6437362CB1 | 3558 | | 1568793H1 (UTRSNOT05) | 1264 | 1467 |

Table 4 (cont.)

| Polynucleotide SEQ ID NO: | Incyte Polynucleotide ID | Sequence Length | Selected Fragment(s) | Sequence Fragments | 5' Position | 3' Position |
|------------------------------|--------------------------------|--------------------|--|-----------------------|-------------|-------------|
| 64 | 6437362CB1 | 3558 | | 987366H1 (LVENNOT03) | 763 | 1088 |
| 64 | 6437362CB1 | 3558 | | 95689540 CD | 678 | 3298 |
| 64 | 6437362CB1 | 3558 | | 4567664F6 (HELATXT01) | 1 | 500 |
| 64 | 6437362CB1 | 3558 | | 1988667R6 (LUNGAST01) | 2160 | 2709 |
| 64 | 6437362CB1 | 3558 | | 4980789H1 (HELATXT04) | 981 | 1248 |
| 64 | 6437362CB1 | 3558 | | 6437362H1 (LUNGNON07) | 432 | 1052 |
| 64 | 6437362CB1 | 3558 | | 2903936F6 (DRGCNOT01) | 1760 | 2328 |
| 64 | 6437362CB1 | 3558 | | 3208061H1 (PENCNOT03) | 2706 | 2948 |
| 64 | 6437362CB1 | 3558 | | 3288660F6 (BONRFET01) | 1433 | 2050 |
| 64 | 6437362CB1 | 3558 | | 865171R1 (BRAITUT03) | 2929 | 3483 |
| 65 | 4173970CB1 | 5373 | 3418-5373, 1-186, 1641-2444, 857-1058 | 829704R1 (PROSTUT04) | 1356 | 1950 |
| 65 | 4173970CB1 | 5373 | | 5604442H1 (MONOTXN03) | 2038 | 2310 |
| 65 | 4173970CB1 | 5373 | | 1708630F6 (PROSNOT16) | 3923 | 4539 |
| 65 | 4173970CB1 | 5373 | | 4561514F6 (KERATXT01) | 4397 | 5197 |
| 65 | 4173970CB1 | 5373 | | 1437088F1 (PANCNOT08) | 3691 | 4247 |
| 65 | 4173970CB1 | 5373 | | 1433309R1 (BEPINON01) | 3666 | 4214 |
| 65 | 4173970CB1 | 5373 | | 4167822F6 (PANCNOT21) | 1 | 494 |
| 65 | 4173970CB1 | 5373 | | 1806736F6 (SINTNOT13) | 5043 | 5373 |
| 65 | 4173970CB1 | 5373 | | 3508537F6 (CONCNOT01) | 2520 | 3017 |
| 65 | 4173970CB1 | 5373 | | 4173970F6 (SINTNOT21) | 2115 | 2677 |
| 65 | 4173970CB1 | 5373 | | 2277402R6 (PROSNON01) | 660 | 1229 |
| 65 | 4173970CB1 | 5373 | | 1708630T6 (PROSNOT16) | 4639 | 5347 |
| 65 | 4173970CB1 | 5373 | | 5944958H1 (COLADIT05) | 1710 | 2018 |
| 65 | 4173970CB1 | 5373 | | 1300156F1 (BRSTNOT07) | 1133 | 1687 |
| 65 | 4173970CB1 | 5373 | | 92737563 | 1885 | 2171 |
| 65 | 4173970CB1 | 5373 | | 209752R1 (SPLNNOT02) | 3040 | 3690 |
| 65 | 4173970CB1 | 5373 | | 287603R1 (EOSIHET02) | 210 | 966 |
| 65 | 4173970CB1 | 5373 | | 3091106H1 (BRSTNOT19) | 4284 | 4560 |
| 65 | 4173970CB1 | 5373 | | 516280R6 (MMLRIDT01) | 2996 | 3438 |
| 66 | 2772751CB1 | 4333 | 2456-3205, 100-160, 1-23, 1459-1957, 3695-4333 | 70475866V1 | 507 | 1008 |
| 66 | 2772751CB1 | 4333 | | 70472414V1 | 2045 | 2690 |

Table 4 (cont.)

| Polynucleotide SEQ ID NO: | Incyte Polynucleotide ID | Sequence Length | Selected Fragment(s) | Sequence Fragments | 5' Position | 3' Position |
|------------------------------|--------------------------------|--------------------|-------------------------|-----------------------|-------------|-------------|
| 66 | 2772751CB1 | 4333 | | 3402774H1 (ESOGNOT03) | 1 | 256 |
| 66 | 2772751CB1 | 4333 | | 70475304V1 | 1481 | 2093 |
| 66 | 2772751CB1 | 4333 | | 70475403V1 | 976 | 1554 |
| 66 | 2772751CB1 | 4333 | | 70470913V1 | 1545 | 2109 |
| 66 | 2772751CB1 | 4333 | | 70747026V1 | 3561 | 4054 |
| 66 | 2772751CB1 | 4333 | | g3923880 | 319 | 678 |
| 66 | 2772751CB1 | 4333 | | 1298255T6 (BRSTNOT07) | 3001 | 3690 |
| 66 | 2772751CB1 | 4333 | | 70472159V1 | 1049 | 1556 |
| 66 | 2772751CB1 | 4333 | | 70472656V1 | 2219 | 2925 |
| 66 | 2772751CB1 | 4333 | | g5340324 | 24 | 467 |
| 66 | 2772751CB1 | 4333 | | 6849173H1 (KIDNTMN03) | 3801 | 4333 |
| 66 | 2772751CB1 | 4333 | | 6221536U1 | 2737 | 3452 |
| 67 | 2793768CB1 | 2213 | 2186-2213, 1066-1156 | 70843048V1 | 1646 | 2213 |
| 67 | 2793768CB1 | 2213 | | 2026465R6 (KERANOT02) | 709 | 1268 |
| 67 | 2793768CB1 | 2213 | | 7712268H1 (TESTTUE02) | 1621 | 2213 |
| 67 | 2793768CB1 | 2213 | | g1958420 | 1 | 421 |
| 67 | 2793768CB1 | 2213 | | 6584157H1 (ESOGTMC01) | 984 | 1576 |
| 67 | 2793768CB1 | 2213 | | 2793768F6 (COLNTUT16) | 131 | 777 |
| 67 | 2793768CB1 | 2213 | | 71279716V1 | 1510 | 2183 |
| 68 | 3035248CB1 | 1142 | 1-55, 555-605 | 71515027V1 | 585 | 1140 |
| 68 | 3035248CB1 | 1142 | | 71486327V1 | 402 | 787 |
| 68 | 3035248CB1 | 1142 | | 71514455V1 | 1 | 576 |

Table 5

| Polynucleotide SEQ ID NO: | IncYTE Project ID | Representative Library |
|------------------------------|----------------------|------------------------|
| 35 | 1889577CB1 | PROSTUT12 |
| 36 | 2427982CB1 | DRGNOT01 |
| 37 | 2470833CB1 | LUNGAST01 |
| 38 | 2080579CB1 | UTRSNOT08 |
| 39 | 2156553CB1 | THYMFET03 |
| 40 | 2182855CB1 | SCORNOT04 |
| 41 | 2242106CB1 | COLNPOT01 |
| 42 | 2726877CB1 | LUNGNOT34 |
| 43 | 2738233CB1 | MENITUT03 |
| 44 | 1833116CB1 | THYRNOT03 |
| 45 | 001799CB1 | BRSTTUT02 |
| 46 | 119814CB1 | MUSCNOT01 |
| 47 | 1295420CB1 | BRAITUT12 |
| 48 | 1309364CB1 | TYLMUNT01 |
| 49 | 1315267CB1 | BLADTUT02 |
| 50 | 1403289CB1 | LATRTUT02 |
| 51 | 1607607CB1 | BRAIDIT01 |
| 52 | 1660025CB1 | BRAWNOT01 |
| 53 | 1796836CB1 | BRSTNOT05 |
| 54 | 2880670CB1 | OVARTUT01 |
| 55 | 2913976CB1 | ENDCNOT04 |
| 56 | 3092084CB1 | HEAANOT01 |
| 57 | 3882482CB1 | SPLNNOT11 |
| 58 | 4933451CB1 | BRSTTUT20 |
| 59 | 5043904CB1 | PLACFER01 |
| 60 | 5202390CB1 | TESTTUT02 |
| 61 | 5526375CB1 | KIDNFET02 |
| 62 | 5677408CB1 | ADRENOT03 |
| 63 | 5982278CB1 | SPLNFET02 |
| 64 | 6437362CB1 | BRAINOT23 |
| 65 | 4173970CB1 | BRSTNOT07 |
| 66 | 2772751CB1 | BRSTNOT07 |
| 67 | 2793768CB1 | UTRSNOT12 |
| 68 | 3035248CB1 | TYLMNOT05 |

Table 6

| Library | Vector | Library Description |
|-----------|---------|---|
| ADRENOT03 | PSPORT1 | Library was constructed using RNA isolated from the adrenal tissue of a 17-year-old Caucasian male, who died from cerebral anoxia. |
| BLADTUT02 | pINCY | Library was constructed using RNA isolated from bladder tumor tissue removed from an 80-year-old Caucasian female during a radical cystectomy and lymph node excision. Pathology indicated grade 3 invasive transitional cell carcinoma. Family history included acute renal failure, osteoarthritis, and atherosclerosis. |
| BRAIDIT01 | pINCY | Library was constructed using RNA isolated from diseased brain tissue. Patient history included multiple sclerosis, type II lesion. |
| BRAINOT23 | pINCY | Library was constructed using RNA isolated from right temporal lobe tissue removed from a 45-year-old Black male during a brain lobectomy. Pathology for the associated tumor tissue indicated dysembryoplastic neuroepithelial tumor of the right temporal lobe. The right temporal region dura was consistent with calcifying pseudotumor of the neuraxis. The patient presented with convulsive intractable epilepsy, partial epilepsy, and memory disturbance. Patient history included obesity, meningitis, backache, unspecified sleep apnea, acute stressreaction, acquired knee deformity, and chronic sinusitis. Family history included obesity, benign hypertension, cirrhosis of the liver, alcohol abuse, hyperlipidemia, cerebrovascular disease, and type II diabetes. |
| BRAITUT12 | pINCY | Library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 40-year-old Caucasian female during excision of a cerebral meningial lesion. Pathology indicated grade 4 gemistocytic astrocytoma. |
| BRAWNOT01 | pINCY | Library was constructed using RNA isolated from dentate nucleus tissue removed from the brain of a 35-year-old Caucasian male who died from cardiac failure. Pathology indicated moderate leptomeningeal fibrosis and multiple microinfarctions of the cerebral neocortex. Grossly, the brain regions examined and cranial nerves were unremarkable, showing no evidence of atrophy. No atherosclerosis of the major vessels was noted. Microscopically, the cerebral hemisphere revealed moderate fibrosis of the leptomeninges with focal calcifications. There was evidence of shrunken and slightly eosinophilic pyramidal neurons throughout the cerebral hemispheres. There were also multiple small microscopic areas of cavitation with surrounding gliosis scattered throughout the cerebral cortex. Special stains with Bielschowsky silver, Klüver-Barrera, and Congo Red revealed no evidence of neurofibrillary tangles or diffuse anorectic amyloid plaques, demyelination, and cerebral amyloid angiopathy, respectively. Patient history included dilated cardiomyopathy, congestive heart failure, cardiomegaly, and an enlarged spleen and liver. Patient medications included simethicone, Lasix, Digoxin, Colace, Zantac, captopril, and Vasotec. |
| BRSTNOT05 | PSPORT1 | Library was constructed using RNA isolated from breast tissue removed from a 58-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology for the associated tumor tissue indicated multicentric invasive grade 4 lobular carcinoma. Patient history included skin cancer, rheumatic heart disease, osteoarthritis, and tuberculosis. Family history included cerebrovascular and cardiovascular disease, breast and prostate cancer, and type I diabetes. |

Table 6 (cont.)

| Library | Vector | Library Description |
|-----------|---------|---|
| BRSTNOT07 | pINCY | Library was constructed using RNA isolated from diseased breast tissue removed from a 43-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated mildly proliferative fibrocystic changes with epithelial hyperplasia, papillomatosis, and duct ectasia. Pathology for the associated tumor tissue indicated invasive grade 4, nuclear grade 3 mammary adenocarcinoma with extensive comedo necrosis. Family history included epilepsy, cardiovascular disease, and type II diabetes. |
| BRSTNOT07 | pINCY | Library was constructed using RNA isolated from diseased breast tissue removed from a 43-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated mildly proliferative fibrocystic changes with epithelial hyperplasia, papillomatosis, and duct ectasia. Pathology for the associated tumor tissue indicated invasive grade 4, nuclear grade 3 mammary adenocarcinoma with extensive comedo necrosis. Family history included epilepsy, cardiovascular disease, and type II diabetes. |
| BRSTTUT02 | PSPORT1 | Library was constructed using RNA isolated from breast tumor tissue removed from a 54-year-old Caucasian female during a bilateral radical mastectomy with reconstruction. Pathology indicated residual invasive grade 3 mammary ductal adenocarcinoma. The remaining breast parenchyma exhibited proliferative fibrocystic changes without atypia. One of 10 axillary lymph nodes had metastatic tumor as a microscopic intranodal focus. Patient history included kidney infection and condyloma acuminatum. Family history included benign hypertension, hyperlipidemia, and a malignant colon neoplasm. |
| BRSTTUT20 | pINCY | Library was constructed using RNA isolated from left breast tumor tissue removed from a 66-year-old Black female during a unilateral extended simple mastectomy and fine needle breast biopsy. Pathology indicated invasive grade 4, nuclear grade 3 adenocarcinoma ductal type, diffusely replacing the left breast. The skin, nipple and fascia were all involved, including the deep surgical margin. Extensive angiolymphatic invasion was identified, including superficial dermal lymphatics. Metastatic grade 4 adenocarcinoma completely replaced 6 lymph nodes with extranodal extension. Multiple low axillary lymph nodes tissue were positive for metastatic mammary carcinoma. Left chest wall biopsy indicated metastatic grade 4 adenocarcinoma. Prior left breast biopsy indicated metastatic grade 4, nuclear grade 3, metastatic mammary carcinoma. The patient presented with malaise and fatigue. Patient history included secondary malignant neoplasm of the liver, secondary malignant neoplasm of the brain/spine, deficiency anemia, type II diabetes, chronic renal failure, and normal delivery. Patient medications included two cycles of cyclophosphamide/epirubicin and 5-Fluorouracil in November 1995. Family history included benign hypertension, type II diabetes, hyperlipidemia, and depressive disorder in the mother. |
| COLNPOT01 | pINCY | Library was constructed using RNA isolated from colon polyp tissue removed from a 40-year-old Caucasian female during a total colectomy. Pathology indicated an inflammatory pseudopolyp; this tissue was associated with a focally invasive grade 2 adenocarcinoma and multiple tubovillous adenomas. Patient history included a benign neoplasm of the bowel. |

Table 6 (cont.)

| Library | Vector | Library Description |
|-----------|-------------|--|
| DRGCNOT01 | pINCY | Library was constructed using RNA isolated from dorsal root ganglion tissue removed from the cervical spine of a 32-year-old Caucasian male who died from acute pulmonary edema and bronchopneumonia, bilateral pleural and pericardial effusions, and malignant lymphoma (natural killer cell type). Patient history included probable cytomegalovirus, infection, hepatic congestion and steatosis, splenomegaly, hemorrhagic cystitis, thyroid hemorrhage, and Bell's palsy. Surgeries included colonoscopy, large intestine biopsy, adenotonsillectomy, and nasopharyngeal endoscopy and biopsy; treatment included radiation therapy. |
| ENDCNOT04 | pINCY | Library was constructed using RNA isolated from coronary artery endothelial cell tissue removed from a 3-year-old Caucasian male. |
| HEANOT01 | pINCY | Library was constructed using RNA isolated from right coronary and right circumflex coronary artery tissue removed from the explanted heart of a 46-year-old Caucasian male during a heart transplantation. Patient history included myocardial infarction from total occlusion of the left anterior descending coronary artery, atherosclerotic coronary artery disease, hyperlipidemia, myocardial ischemia, dilated cardiomyopathy, left ventricular dysfunction, and tobacco abuse. Previous surgeries included cardiac catheterization. Family history included atherosclerotic coronary artery disease. |
| KIDNFET02 | pINCY | Library was constructed using RNA isolated from kidney tissue removed from a Caucasian male fetus, who was stillborn with a hypoplastic left heart and died at 23 weeks' gestation. |
| LATRTUT02 | pINCY | Library was constructed using RNA isolated from a myxoma removed from the left atrium of a 43-year-old Caucasian male during annuloplasty. Pathology indicated atrial myxoma. Patient history included pulmonary insufficiency, acute myocardial infarction, atherosclerotic coronary artery disease, hyperlipidemia, and tobacco use. Family history included benign hypertension, acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes. |
| LUNGAST01 | PSPORT1 | Library was constructed using RNA isolated from the lung tissue of a 17-year-old Caucasian male, who died from head trauma. Patient history included asthma. |
| LUNGNOT34 | pINCY | Library was constructed using RNA isolated from lung tissue removed from a 12-year-old Caucasian male. |
| MENITUT03 | pINCY | Library was constructed using RNA isolated from brain meningioma tissue removed from a 35-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology indicated a benign neoplasm in the right cerebellopontine angle of the brain. Patient history included hypothyroidism. Family history included myocardial infarction and breast cancer. |
| MUSCNOT01 | PBLUESCRIPT | Library was constructed at Stratagene (STR937209), using RNA isolated from the skeletal muscle tissue of a patient with malignant hyperthermia. |

Table 6 (cont.)

| Library | Vector | Library Description |
|------------|---------|--|
| OVARTUT01 | PSPORT1 | Library was constructed using RNA isolated from ovarian tumor tissue removed from a 43-year-old Caucasian female during removal of the fallopian tubes and ovaries. Pathology indicated grade 2 mucinous cystadenocarcinoma involving the entire left ovary. Patient history included mitral valve disorder, pneumonia, and viral hepatitis. Family history included atherosclerotic coronary artery disease, pancreatic cancer, stress reaction, cerebrovascular disease, breast cancer, and uterine cancer. |
| PLACFER01 | pINCY | The library was constructed using RNA isolated from placental tissue removed from a Caucasian fetus, who died after 16 weeks' gestation from fetal demise and hydrocephalus. Patient history included umbilical cord wrapped around the head (3 times) and the shoulders (1 time). Serology was positive for anti-CMV. Family history included multiple pregnancies and live births, and an abortion. |
| PROSTUT12 | pINCY | Library was constructed using RNA isolated from prostate tumor tissue removed from a 65-year-old Caucasian male during a radical prostatectomy. Pathology indicated an adenocarcinoma (Gleason grade 2+2). Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA). |
| SCORNOT04 | pINCY | Library was constructed using RNA isolated from cervical spinal cord tissue removed from a 32-year-old Caucasian male who died from acute pulmonary edema and bronchopneumonia, bilateral pleural and pericardial effusions, and malignant lymphoma (natural killer cell type). Patient history included probable cytomegalovirus, infection, hepatic congestion and steatosis, splenomegaly, hemorrhagic cystitis, thyroid hemorrhage, and Bell's palsy. Surgeries included colonoscopy, large intestine biopsy, adenotonsillectomy, and nasopharyngeal endoscopy and biopsy; treatment included radiation therapy. |
| SPLNFET02 | pINCY | Library was constructed using RNA isolated from spleen tissue removed from a Caucasian male fetus, who died at 23 weeks' gestation. |
| SPLNNOT11 | pINCY | Library was constructed using RNA isolated from diseased spleen tissue removed from a 14-year-old Asian male during a total splenectomy. Pathology indicated changes consistent with idiopathic thrombocytopenic purpura. The patient presented with bruising. Patient medications included Vincristine. |
| TESTTUT02 | pINCY | Library was constructed using RNA isolated from testicular tumor removed from a 31-year-old Caucasian male during unilateral orchiectomy. Pathology indicated embryonal carcinoma. |
| THYMFET03 | pINCY | Library was constructed using RNA isolated from thymus tissue removed from a Caucasian male fetus. |
| THYRNOT03 | pINCY | Library was constructed using RNA isolated from thyroid tissue removed from the left thyroid of a 28-year-old Caucasian female during a complete thyroidectomy. Pathology indicated a small nodule of adenomatous hyperplasia present in the left thyroid. Pathology for the associated tumor tissue indicated dominant follicular adenoma, forming a well-encapsulated mass in the left thyroid. |
| TYLYMNOT05 | pINCY | Library was constructed RNA isolated from nonactivated Th2 cells. These cells were differentiated from umbilical cord CD4 T cells with IL-4 in the presence of anti-IL-12 antibodies and B7-transfected COS cells. |

Table 6 (cont.)

| Library | Vector | Library Description |
|-----------|--------|---|
| TYMUNT01 | pINCY | Library was constructed using RNA isolated from resting allogenic T-lymphocyte tissue removed from an adult (40-50-year-old) Caucasian male. |
| UTRSNOT08 | pINCY | Library was constructed using RNA isolated from uterine tissue removed from a 35-year-old Caucasian female during a vaginal hysterectomy with dilation and curettage. Pathology indicated that the endometrium was secretory phase with a benign endometrial polyp 1 cm in diameter. The cervix showed mild chronic cervicitis. Family history included atherosclerotic coronary artery disease and type II diabetes. |
| UTRSNOT12 | pINCY | Library was constructed using RNA isolated from uterine myometrial tissue removed from a 41-year-old Caucasian female during a vaginal hysterectomy with dilation and curettage. The endometrium was secretory and contained fragments of endometrial polyps. Benign endo- and ectocervical mucosa were identified in the endocervix. Pathology for the associated tumor tissue indicated uterine leiomyoma. Patient history included ventral hernia and a benign ovarian neoplasm. |

Table 7

| Program | Description | Reference | Parameter Threshold |
|-------------------|---|--|---|
| ABI FACTURA | A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences. | Applied Biosystems, Foster City, CA. | |
| ABI/PARACEL FDF | A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences. | Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA. | Mismatch <50% |
| ABI AutoAssembler | A program that assembles nucleic acid sequences. | Applied Biosystems, Foster City, CA. | |
| BLAST | A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx. | Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402. | ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less |
| FASTA | A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and ssearch. | Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489. | ESTs: fasta E value=1.06E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less Full Length sequences: fastx score=100 or greater |
| BLIMPS | A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions. | Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424. | Probability value= 1.0E-3 or less |
| HMMER | An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM. | Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350. | PFAM hits: Probability value= 1.0E-3 or less Signal peptide hits: Score= 0 or greater |

Table 7 (cont.)

| Program | Description | Reference | Parameter Threshold |
|-------------|---|--|---|
| ProfileScan | An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite. | Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221. | Normalized quality score > GCG-specified "HIGH" value for that particular Prosite motif. Generally, score = 1.4-2.1. |
| Phred | A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability. | Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194. | |
| Phrap | A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences. | Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA. | Score = 120 or greater; Match length = 56 or greater |
| Consed | A graphical tool for viewing and editing Phrap assemblies. | Gordon, D. et al. (1998) Genome Res. 8:195-202. | |
| SPScan | A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides. | Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439. | Score = 3.5 or greater |
| TMAP | A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation. | Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371. | |
| TMHMMER | A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation. | Sonnhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182. | |
| Motifs | A program that searches amino acid sequences for patterns that matched those defined in Prosite. | Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI. | |

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:
 - a) a polypeptide comprising an amino acid sequence selected from the group consisting of
5 SEQ ID NO:1-34,
 - b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-34,
 - c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34, and
10 d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34.
2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-
15 34.
3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide encoding a polypeptide of claim 2.
- 20 5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:35-68.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
- 25 7. A cell transformed with a recombinant polynucleotide of claim 6.
8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
- 30 9. A method for producing a polypeptide of claim 1, the method comprising:
 - a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and
35 b) recovering the polypeptide so expressed.

10. An isolated antibody which specifically binds to a polypeptide of claim 1.
11. An isolated polynucleotide selected from the group consisting of:
- a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting
 - 5 of SEQ ID NO:35-68,
 - b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:35-68,
 - c) a polynucleotide complementary to a polynucleotide of a),
 - d) a polynucleotide complementary to a polynucleotide of b), and
 - 10 e) an RNA equivalent of a)-d).
12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.
13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:
- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization
 - 20 complex is formed between said probe and said target polynucleotide or fragments thereof, and
 - b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.
14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.
15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:
- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
 - 30 b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.
16. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

35

17. A composition of claim 16, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-34.

18. A method for treating a disease or condition associated with decreased expression of functional CYSKP, comprising administering to a patient in need of such treatment the composition of claim 16.

19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

10 a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

 b) detecting agonist activity in the sample.

20. A composition comprising an agonist compound identified by a method of claim 19 and a pharmaceutically acceptable excipient.

15

21. A method for treating a disease or condition associated with decreased expression of functional CYSKP, comprising administering to a patient in need of such treatment a composition of claim 20.

22. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

20 a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

 b) detecting antagonist activity in the sample.

23. A composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.

25

24. A method for treating a disease or condition associated with overexpression of functional CYSKP, comprising administering to a patient in need of such treatment a composition of claim 23.

30

25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:

 a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and

35 b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a

compound that specifically binds to the polypeptide of claim 1.

26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:

- 5 a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in
- 10 the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

27. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method

15 comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,
- b) detecting altered expression of the target polynucleotide, and
- c) comparing the expression of the target polynucleotide in the presence of varying amounts of
- 20 the compound and in the absence of the compound.

28. A method for assessing toxicity of a test compound, said method comprising:

- a) treating a biological sample containing nucleic acids with the test compound;
- b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at
- 25 least 20 contiguous nucleotides of a polynucleotide of claim 11 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 11 or fragment thereof;
- c) quantifying the amount of hybridization complex; and
- 30 d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

29. A diagnostic test for a condition or disease associated with the expression of CYSKP in a biological sample comprising the steps of:

- a) combining the biological sample with an antibody of claim 10, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex; and
- 5 b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

30. The antibody of claim 10, wherein the antibody is:

- a) a chimeric antibody,
- 10 b) a single chain antibody,
- c) a Fab fragment,
- d) a F(ab')₂ fragment, or
- e) a humanized antibody.

15 31. A composition comprising an antibody of claim 10 and an acceptable excipient.

32. A method of diagnosing a condition or disease associated with the expression of CYSKP in a subject, comprising administering to said subject an effective amount of the composition of claim 31.

20

33. A composition of claim 31, wherein the antibody is labeled.

34. A method of diagnosing a condition or disease associated with the expression of CYSKP in a subject, comprising administering to said subject an effective amount of the composition of claim 33.

25

35. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 10 comprising:

- a) immunizing an animal with a polypeptide having an amino acid sequence selected from
- 30 the group consisting of SEQ ID NO:1-34, or an immunogenic fragment thereof, under conditions to elicit an antibody response;
- b) isolating antibodies from said animal; and
- c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the

group consisting of SEQ ID NO:1-34.

36. An antibody produced by a method of claim 35.

5 37. A composition comprising the antibody of claim 36 and a suitable carrier.

38. A method of making a monoclonal antibody with the specificity of the antibody of claim
10 comprising:

- 10 a) immunizing an animal with a polypeptide having an amino acid sequence selected from
the group consisting of SEQ ID NO:1-34, or an immunogenic fragment thereof, under conditions to
elicit an antibody response;
- b) isolating antibody producing cells from the animal;
- c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-
producing hybridoma cells;
- 15 d) culturing the hybridoma cells; and
- e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide
having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34.

39. A monoclonal antibody produced by a method of claim 38.

20

40. A composition comprising the antibody of claim 39 and a suitable carrier.

41. The antibody of claim 10, wherein the antibody is produced by screening a Fab
expression library.

25

42. The antibody of claim 10, wherein the antibody is produced by screening a recombinant
immunoglobulin library.

43. A method for detecting a polypeptide having an amino acid sequence selected from the
30 group consisting of SEQ ID NO:1-34 in a sample, comprising the steps of:

- a) incubating the antibody of claim 10 with a sample under conditions to allow specific
binding of the antibody and the polypeptide; and
- b) detecting specific binding, wherein specific binding indicates the presence of a
polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34 in

the sample.

44. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34 from a sample, the method comprising:

- 5 a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and
- b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-34.

10 45. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.

46. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

47. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

15

48. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

49. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

20

50. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

51. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.

52. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

25

53. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.

54. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.

30

55. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.

56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.

57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.

58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.
59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.
- 5 60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.
61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.
- 10 62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.
63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.
64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20.
- 15 65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:21.
66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:22.
- 20 67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:23.
68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:24.
69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:25.
- 25 70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:26.
71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:27.
- 30 72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:28.
73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:29.
74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:30.

75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:31.
76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:32.
- 5 77. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:33.
78. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:34.
- 10 79. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:35.
80. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:36.
- 15 81. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:37.
82. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:38.
- 20 83. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:39.
84. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:40.
- 25 85. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:41.
- 30 86. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:42.
87. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID

NO:43.

88. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:44.

5

89. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:45.

90. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
10 NO:46.

91. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:47.

15 92. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:48.

93. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:49.

20

94. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:50.

25 95. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:51.

96. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:52.

30 97. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:53.

98. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:54.

99. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:55.

5 100. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:56.

101. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:57.

10 102. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:58.

15 103. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:59.

104. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:60.

20 105. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:61.

106. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:62.

25 107. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:63.

30 108. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:64.

109. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:65.

110. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:66.

111. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
5 NO:67.

112. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:68.

<110> INCYTE GENOMICS, INC.
 YUE, Henry
 TANG, Y. Tom
 AU-YOUNG, Janice
 LU, Dyung Aina M.
 BAUGHN, Mariah R.
 HILLMAN, Jennifer L.
 AZIMZAI, Yalda
 LAL, Preeti
 YAO, Monique G.
 BANDMAN, Olga
 BURFORD, Neil
 BATRA, Sajeev
 KEARNEY, Liam
 POLICKY, Jennifer L.

<120> CYTOSKELETON-ASSOCIATED PROTEINS

<130> PF-0772 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/201,960; 60/202,729; 60/209,705; 60/210,149; 60/213,215

<151> 2000-05-05; 2000-05-08; 2000-06-05; 2000-06-07; 2000-06-21

<160> 68

<170> PERL Program

<210> 1

<211> 622

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1889577CD1

<400> 1

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ala | Met | Met | Val | Phe | Pro | Arg | Glu | Glu | Lys | Leu | Ser | Gln | Asp |
| 1 | | | | 5 | | | | | 10 | | | | | 15 |
| Glu | Ile | Val | Leu | Gly | Thr | Lys | Ala | Val | Ile | Gln | Gly | Leu | Glu | Thr |
| | | | | 20 | | | | | 25 | | | | | 30 |
| Leu | Arg | Gly | Glu | His | Arg | Ala | Leu | Leu | Ala | Pro | Leu | Val | Ala | Pro |
| | | | | 35 | | | | | 40 | | | | | 45 |
| Glu | Ala | Gly | Glu | Pro | Glu | Pro | Gly | Ser | Gln | Glu | Arg | Cys | Ile | Leu |
| | | | | 50 | | | | | 55 | | | | | 60 |
| Leu | Arg | Arg | Ser | Leu | Glu | Ala | Ile | Glu | Leu | Gly | Leu | Gly | Glu | Ala |
| | | | | 65 | | | | | 70 | | | | | 75 |
| Gln | Val | Ile | Leu | Ala | Leu | Ser | Ser | His | Leu | Gly | Ala | Val | Glu | Ser |
| | | | | 80 | | | | | 85 | | | | | 90 |
| Glu | Lys | Gln | Lys | Leu | Arg | Ala | Gln | Val | Arg | Arg | Leu | Val | Gln | Glu |
| | | | | 95 | | | | | 100 | | | | | 105 |
| Asn | Gln | Trp | Leu | Arg | Glu | Glu | Leu | Ala | Gly | Thr | Gln | Gln | Lys | Leu |
| | | | | 110 | | | | | 115 | | | | | 120 |
| Gln | Arg | Ser | Glu | Gln | Ala | Val | Ala | Gln | Leu | Glu | Glu | Glu | Lys | Gln |
| | | | | 125 | | | | | 130 | | | | | 135 |
| His | Leu | Leu | Phe | Met | Ser | Gln | Ile | Arg | Lys | Leu | Asp | Glu | Asp | Ala |
| | | | | 140 | | | | | 145 | | | | | 150 |
| Ser | Pro | Asn | Glu | Glu | Lys | Gly | Asp | Val | Pro | Lys | Asp | Thr | Leu | Asp |
| | | | | 155 | | | | | 160 | | | | | 165 |
| Asp | Leu | Phe | Pro | Asn | Glu | Asp | Glu | Gln | Ser | Pro | Ala | Pro | Ser | Pro |
| | | | | 170 | | | | | 175 | | | | | 180 |
| Gly | Gly | Gly | Asp | Val | Ser | Gly | Gln | His | Gly | Gly | Tyr | Glu | Ile | Pro |
| | | | | 185 | | | | | 190 | | | | | 195 |
| Ala | Arg | Leu | Arg | Thr | Leu | His | Asn | Leu | Val | Ile | Gln | Tyr | Ala | Ser |

| | | | | | |
|-----------------|---------------------|---------------------|-----|--|-----|
| | 200 | | 205 | | 210 |
| Gln Gly Arg Tyr | Glu Val Ala Val Pro | Leu Cys Lys Gln Ala | Leu | | |
| | 215 | | 220 | | 225 |
| Glu Asp Leu Glu | Lys Thr Ser Gly His | Asp His Pro Asp Val | Ala | | |
| | 230 | | 235 | | 240 |
| Thr Met Leu Asn | Ile Leu Ala Leu Val | Tyr Arg Asp Gln Asn | Lys | | |
| | 245 | | 250 | | 255 |
| Tyr Lys Glu Ala | Ala His Leu Leu Asn | Asp Ala Leu Ala Ile | Arg | | |
| | 260 | | 265 | | 270 |
| Glu Lys Thr Leu | Gly Lys Asp His Pro | Ala Val Ala Ala Thr | Leu | | |
| | 275 | | 280 | | 285 |
| Asn Asn Leu Ala | Val Leu Tyr Gly Lys | Arg Gly Lys Tyr Lys | Glu | | |
| | 290 | | 295 | | 300 |
| Ala Glu Pro Leu | Cys Lys Arg Ala Leu | Glu Ile Arg Glu Lys | Val | | |
| | 305 | | 310 | | 315 |
| Leu Gly Lys Phe | His Pro Asp Val Ala | Lys Gln Leu Ser Asn | Leu | | |
| | 320 | | 325 | | 330 |
| Ala Leu Leu Cys | Gln Asn Gln Gly Lys | Ala Glu Glu Val Glu | Tyr | | |
| | 335 | | 340 | | 345 |
| Tyr Tyr Arg Arg | Ala Leu Glu Ile Tyr | Ala Thr Arg Leu Gly | Pro | | |
| | 350 | | 355 | | 360 |
| Asp Asp Pro Asn | Val Ala Lys Thr Lys | Asn Asn Leu Ala Ser | Cys | | |
| | 365 | | 370 | | 375 |
| Tyr Leu Lys Gln | Gly Lys Tyr Gln Asp | Ala Glu Thr Leu Tyr | Lys | | |
| | 380 | | 385 | | 390 |
| Glu Ile Leu Thr | Arg Ala His Glu Lys | Glu Phe Gly Ser Val | Asn | | |
| | 395 | | 400 | | 405 |
| Gly Asp Asn Lys | Pro Ile Trp Met His | Ala Glu Glu Arg Glu | Glu | | |
| | 410 | | 415 | | 420 |
| Ser Lys Asp Lys | Arg Arg Asp Ser Ala | Pro Tyr Gly Glu Tyr | Gly | | |
| | 425 | | 430 | | 435 |
| Ser Trp Tyr Lys | Ala Cys Lys Val Asp | Ser Pro Thr Val Asn | Thr | | |
| | 440 | | 445 | | 450 |
| Thr Leu Arg Thr | Leu Gly Ala Leu Tyr | Arg Arg Gln Gly Lys | Leu | | |
| | 455 | | 460 | | 465 |
| Glu Ala Ala His | Thr Leu Glu Asp Cys | Ala Ser Arg Asn Arg | Lys | | |
| | 470 | | 475 | | 480 |
| Gln Gly Leu Asp | Pro Ala Ser Gln Thr | Lys Val Val Glu Leu | Leu | | |
| | 485 | | 490 | | 495 |
| Lys Asp Gly Ser | Gly Arg Arg Gly Asp | Pro Arg Ser Ser Arg | Asp | | |
| | 500 | | 505 | | 510 |
| Met Ala Gly Gly | Ala Gly Pro Arg Ser | Glu Ser Asp Leu Glu | Asp | | |
| | 515 | | 520 | | 525 |
| Val Gly Pro Thr | Ala Glu Trp Asn Gly | Asp Gly Ser Gly Ser | Leu | | |
| | 530 | | 535 | | 540 |
| Arg Arg Ser Gly | Ser Phe Gly Lys Leu | Arg Asp Ala Leu Arg | Arg | | |
| | 545 | | 550 | | 555 |
| Ser Ser Glu Met | Leu Val Lys Lys Leu | Gln Gly Gly Thr Pro | Arg | | |
| | 560 | | 565 | | 570 |
| Glu Pro Pro Asn | Pro Arg Met Lys Arg | Ala Ser Ser Leu Asn | Phe | | |
| | 575 | | 580 | | 585 |
| Leu Asn Lys Ser | Val Glu Glu Pro Thr | Gln Pro Gly Gly Thr | Gly | | |
| | 590 | | 595 | | 600 |
| Leu Ser Asp Ser | Arg Thr Leu Ser Ser | Ser Ser Met Asp Leu | Ser | | |
| | 605 | | 610 | | 615 |
| Arg Arg Ser Ser | Leu Val Gly | | | | |
| | 620 | | | | |

<210> 2

<211> 190

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2427982CD1

<400> 2

```

Met Ala Lys Ala Thr Thr Ile Lys Glu Ala Leu Ala Arg Trp Glu
 1      5      10      15
Glu Lys Thr Gly Gln Arg Pro Ser Glu Ala Lys Glu Ile Lys Leu
 20      25      30
Tyr Ala Gln Ile Pro Ile Glu Lys Met Asp Ala Ser Leu Ser
 35      40      45
Met Leu Ala Asn Cys Glu Lys Leu Ser Leu Ser Thr Asn Cys Ile
 50      55      60
Glu Lys Ile Ala Asn Leu Asn Gly Leu Lys Asn Leu Arg Ile Leu
 65      70      75
Ser Leu Gly Arg Asn Asn Ile Lys Asn Leu Asn Gly Leu Glu Ala
 80      85      90
Val Gly Asp Thr Leu Glu Glu Leu Trp Ile Ser Tyr Asn Phe Ile
 95      100     105
Glu Lys Leu Lys Gly Ile His Ile Met Lys Lys Leu Lys Ile Leu
110     115     120
Tyr Met Ser Asn Asn Leu Val Lys Asp Trp Ala Glu Phe Val Lys
125     130     135
Leu Ala Glu Leu Pro Cys Leu Glu Asp Leu Val Phe Val Gly Asn
140     145     150
Pro Leu Glu Glu Lys His Ser Ala Glu Asn Asn Trp Ile Glu Glu
155     160     165
Ala Thr Lys Arg Val Pro Lys Leu Lys Lys Leu Asp Gly Thr Pro
170     175     180
Val Ile Lys Gly Asp Glu Glu Glu Asp Asn
185     190

```

<210> 3

<211> 331

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2470833CD1

<400> 3

```

Met Glu Pro Glu Phe Leu Tyr Asp Leu Leu Gln Leu Pro Lys Gly
 1      5      10      15
Val Glu Pro Pro Ala Glu Glu Glu Leu Ser Lys Gly Gly Lys Lys
 20      25      30
Lys Tyr Leu Pro Pro Thr Ser Arg Lys Asp Pro Lys Phe Glu Glu
 35      40      45
Leu Gln Lys Val Leu Met Glu Trp Ile Asn Ala Thr Leu Leu Pro
 50      55      60
Glu His Ile Val Val Arg Ser Leu Glu Glu Asp Met Phe Asp Gly
 65      70      75
Leu Ile Leu His His Leu Phe Gln Arg Leu Ala Ala Leu Lys Leu
 80      85      90
Glu Ala Glu Asp Ile Ala Leu Thr Ala Thr Ser Gln Lys His Lys
 95      100     105
Leu Thr Val Val Leu Glu Ala Val Asn Arg Ser Leu Gln Leu Glu
110     115     120
Glu Trp Gln Ala Lys Trp Ser Val Glu Ser Ile Phe Asn Lys Asp
125     130     135
Leu Leu Ser Thr Leu His Leu Leu Val Ala Leu Ala Lys Arg Phe
140     145     150
Gln Pro Asp Leu Ser Leu Pro Thr Asn Val Gln Val Glu Val Ile
155     160     165
Thr Ile Glu Ser Thr Lys Ser Gly Leu Lys Ser Glu Lys Leu Val
170     175     180
Glu Gln Leu Thr Glu Tyr Ser Thr Asp Lys Asp Glu Pro Pro Lys
185     190     195
Asp Val Phe Asp Glu Leu Phe Lys Leu Ala Pro Glu Lys Val Asn
200     205     210
Ala Val Lys Glu Ala Ile Val Asn Phe Val Asn Gln Lys Leu Asp

```

| | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Arg | Leu | Gly | Leu | Ser | Val | Gln | Asn | Leu | Asp | Thr | Gln | Phe | Ala | Asp | 215 | 220 | 225 |
| | | | | | | | | | | | | | | | 230 | 235 | 240 |
| Gly | Val | Ile | Leu | Leu | Leu | Leu | Ile | Gly | Gln | Leu | Glu | Gly | Phe | Phe | 245 | 250 | 255 |
| Leu | His | Leu | Lys | Glu | Phe | Tyr | Leu | Thr | Pro | Asn | Ser | Pro | Ala | Glu | 260 | 265 | 270 |
| Met | Leu | His | Asn | Val | Thr | Leu | Ala | Leu | Glu | Leu | Leu | Lys | Asp | Glu | 275 | 280 | 285 |
| Gly | Leu | Leu | Ser | Cys | Pro | Val | Ser | Pro | Glu | Asp | Ile | Val | Asn | Lys | 290 | 295 | 300 |
| Asp | Ala | Lys | Ser | Thr | Leu | Arg | Val | Leu | Tyr | Gly | Leu | Phe | Cys | Lys | 305 | 310 | 315 |
| His | Thr | Gln | Lys | Ala | His | Arg | Asp | Arg | Thr | Pro | His | Gly | Ala | Pro | 320 | 325 | 330 |

Asn

<210> 4
 <211> 239
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2080579CD1

| | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Met | Glu | Ile | Gln | Met | Asp | Glu | Gly | Gly | Gly | Val | Val | Val | Tyr | 1 | 5 | 10 | 15 |
| Gln | Asp | Asp | Tyr | Cys | Ser | Gly | Ser | Val | Met | Ser | Glu | Arg | Val | Ser | 20 | 25 | 30 | 35 |
| Gly | Leu | Ala | Gly | Ser | Ile | Tyr | Arg | Glu | Phe | Glu | Arg | Leu | Ile | His | 35 | 40 | 45 | 50 |
| Cys | Tyr | Asp | Glu | Glu | Val | Val | Lys | Glu | Leu | Met | Pro | Leu | Val | Val | 50 | 55 | 60 | 65 |
| Asn | Val | Leu | Glu | Asn | Leu | Asp | Ser | Val | Leu | Ser | Glu | Asn | Gln | Glu | 65 | 70 | 75 | 80 |
| His | Glu | Val | Glu | Leu | Glu | Leu | Leu | Arg | Glu | Asp | Asn | Glu | Gln | Leu | 80 | 85 | 90 | 95 |
| Leu | Thr | Gln | Tyr | Glu | Arg | Glu | Lys | Ala | Leu | Arg | Arg | Gln | Ala | Glu | 95 | 100 | 105 | 110 |
| Glu | Lys | Phe | Ile | Glu | Phe | Glu | Asp | Ala | Leu | Glu | Gln | Glu | Lys | Lys | 110 | 115 | 120 | 125 |
| Glu | Leu | Gln | Ile | Gln | Val | Glu | His | Tyr | Glu | Phe | Gln | Thr | Arg | Gln | 125 | 130 | 135 | 140 |
| Leu | Glu | Leu | Lys | Ala | Lys | Asn | Tyr | Ala | Asp | Gln | Ile | Ser | Arg | Leu | 140 | 145 | 150 | 155 |
| Glu | Glu | Arg | Glu | Ser | Glu | Met | Lys | Lys | Glu | Tyr | Asn | Ala | Leu | His | 155 | 160 | 165 | 170 |
| Gln | Arg | His | Thr | Glu | Met | Ile | Gln | Thr | Tyr | Val | Glu | His | Ile | Glu | 170 | 175 | 180 | 185 |
| Arg | Ser | Lys | Met | Gln | Gln | Val | Gly | Gly | Asn | Ser | Gln | Thr | Glu | Ser | 185 | 190 | 195 | 200 |
| Ser | Leu | Pro | Gly | Arg | Arg | Tyr | Ala | Gly | Arg | Gly | Gly | Val | Glu | Val | 200 | 205 | 210 | 215 |
| Arg | Gly | Ala | Arg | Arg | Gly | Gly | Gly | Thr | Gln | Asp | Ala | Ala | His | Ala | 215 | 220 | 225 | 230 |
| Arg | Val | Val | Val | Leu | Val | Met | Ala | Arg | Ala | Leu | Gly | Ser | Gly | | 230 | 235 | | |

<210> 5
 <211> 488
 <212> PRT
 <213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2156553CD1

<400> 5

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asp | Ile | Asp | Lys | Asp | Leu | Glu | Ala | Pro | Leu | Tyr | Leu | Thr | Pro |
| 1 | | | | 5 | | | | | 10 | | | | | 15 |
| Glu | Gly | Trp | Ser | Leu | Phe | Leu | Gln | Arg | Tyr | Tyr | Gln | Val | Val | His |
| | | | | 20 | | | | | 25 | | | | | 30 |
| Glu | Gly | Ala | Glu | Leu | Arg | His | Leu | Asp | Thr | Gln | Val | Gln | Arg | Cys |
| | | | | 35 | | | | | 40 | | | | | 45 |
| Glu | Asp | Ile | Leu | Gln | Gln | Leu | Gln | Ala | Val | Val | Pro | Gln | Ile | Asp |
| | | | | 50 | | | | | 55 | | | | | 60 |
| Met | Glu | Gly | Asp | Arg | Asn | Ile | Trp | Ile | Val | Lys | Pro | Gly | Ala | Lys |
| | | | | 65 | | | | | 70 | | | | | 75 |
| Ser | Arg | Gly | Arg | Gly | Ile | Met | Cys | Met | Asp | His | Leu | Glu | Glu | Met |
| | | | | 80 | | | | | 85 | | | | | 90 |
| Leu | Lys | Leu | Val | Asn | Gly | Asn | Pro | Val | Val | Met | Lys | Asp | Gly | Lys |
| | | | | 95 | | | | | 100 | | | | | 105 |
| Trp | Val | Val | Gln | Lys | Tyr | Ile | Glu | Arg | Pro | Leu | Leu | Ile | Phe | Gly |
| | | | | 110 | | | | | 115 | | | | | 120 |
| Thr | Lys | Phe | Asp | Leu | Arg | Gln | Trp | Phe | Leu | Val | Thr | Asp | Trp | Asn |
| | | | | 125 | | | | | 130 | | | | | 135 |
| Pro | Leu | Thr | Val | Trp | Phe | Tyr | Arg | Asp | Ser | Tyr | Ile | Arg | Phe | Ser |
| | | | | 140 | | | | | 145 | | | | | 150 |
| Thr | Gln | Pro | Phe | Ser | Leu | Lys | Asn | Leu | Asp | Asn | Ser | Val | His | Leu |
| | | | | 155 | | | | | 160 | | | | | 165 |
| Cys | Asn | Asn | Ser | Ile | Gln | Lys | His | Leu | Glu | Asn | Ser | Cys | His | Arg |
| | | | | 170 | | | | | 175 | | | | | 180 |
| His | Pro | Leu | Leu | Pro | Pro | Asp | Asn | Met | Trp | Ser | Ser | Gln | Arg | Phe |
| | | | | 185 | | | | | 190 | | | | | 195 |
| Gln | Ala | His | Leu | Gln | Glu | Met | Gly | Ala | Pro | Asn | Ala | Trp | Ser | Thr |
| | | | | 200 | | | | | 205 | | | | | 210 |
| Ile | Ile | Val | Pro | Gly | Met | Lys | Asp | Ala | Val | Ile | His | Ala | Leu | Gln |
| | | | | 215 | | | | | 220 | | | | | 225 |
| Thr | Ser | Gln | Asp | Thr | Val | Gln | Cys | Arg | Lys | Ala | Ser | Phe | Glu | Leu |
| | | | | 230 | | | | | 235 | | | | | 240 |
| Tyr | Gly | Ala | Asp | Phe | Val | Phe | Gly | Glu | Asp | Phe | Gln | Pro | Trp | Leu |
| | | | | 245 | | | | | 250 | | | | | 255 |
| Ile | Glu | Ile | Asn | Ala | Ser | Pro | Thr | Met | Ala | Pro | Ser | Thr | Ala | Val |
| | | | | 260 | | | | | 265 | | | | | 270 |
| Thr | Ala | Arg | Leu | Cys | Ala | Gly | Val | Gln | Ala | Asp | Thr | Leu | Arg | Val |
| | | | | 275 | | | | | 280 | | | | | 285 |
| Val | Ile | Asp | Arg | Met | Leu | Asp | Arg | Asn | Cys | Asp | Thr | Gly | Ala | Phe |
| | | | | 290 | | | | | 295 | | | | | 300 |
| Glu | Leu | Ile | Tyr | Lys | Gln | Pro | Ala | Val | Glu | Val | Pro | Gln | Tyr | Val |
| | | | | 305 | | | | | 310 | | | | | 315 |
| Gly | Ile | Arg | Leu | Leu | Val | Glu | Gly | Phe | Thr | Ile | Lys | Lys | Pro | Met |
| | | | | 320 | | | | | 325 | | | | | 330 |
| Ala | Met | Cys | His | Arg | Arg | Met | Gly | Val | Arg | Pro | Ala | Val | Pro | Leu |
| | | | | 335 | | | | | 340 | | | | | 345 |
| Leu | Thr | Gln | Arg | Gly | Ser | Gly | Glu | Gly | Lys | Asp | Ser | Gly | Ile | Pro |
| | | | | 350 | | | | | 355 | | | | | 360 |
| Thr | His | Arg | Ser | Ala | Ser | Arg | Lys | Gly | Thr | Gly | Ala | Arg | Ser | Leu |
| | | | | 365 | | | | | 370 | | | | | 375 |
| Gly | His | Ser | Glu | Lys | Pro | Val | Ser | Thr | Ala | Thr | Thr | Ser | Ala | Pro |
| | | | | 380 | | | | | 385 | | | | | 390 |
| Gly | Lys | Gly | Lys | Lys | Gly | Lys | Ala | Lys | Arg | Ala | Thr | Ala | Leu | Val |
| | | | | 395 | | | | | 400 | | | | | 405 |
| Cys | Pro | Asn | Leu | Trp | Glu | Trp | Asp | Ala | Pro | Ser | Thr | Arg | Met | Gly |
| | | | | 410 | | | | | 415 | | | | | 420 |
| Cys | Ile | Phe | Thr | Met | Thr | Phe | Ser | Ser | Gly | Asp | Arg | Gln | Pro | His |
| | | | | 425 | | | | | 430 | | | | | 435 |
| His | Leu | Asn | Arg | Leu | Pro | Leu | Ser | Pro | Lys | Asn | Pro | Gln | Ala | Leu |
| | | | | 440 | | | | | 445 | | | | | 450 |
| Gly | Lys | Thr | Ile | Pro | Pro | Lys | His | Pro | Ser | Val | Pro | Arg | Arg | Phe |
| | | | | 455 | | | | | 460 | | | | | 465 |
| Ile | Pro | Ala | Leu | Gln | Ala | Pro | Pro | Asn | His | Leu | Asp | Gln | Pro | Pro |

His Gln Arg Ala Thr Ser Ser Lys
470
485

475

480

<210> 6
<211> 1190
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2182855CD1

<400> 6
Met Pro Arg Gly Val Phe Gln Gln Leu Ser Asn Leu Val Leu Gln
1 5 10 15
Glu Leu Asn Ala Asn Leu Ser Asn Leu Thr Ser Ala Phe Glu Lys
20 25 30
Ala Thr Ala Glu Lys Ile Lys Cys Gln Gln Glu Ala Asp Ala Thr
35 40 45
Asn Arg Val Ile Leu Leu Ala Asn Arg Leu Val Gly Gly Leu Ala
50 55 60
Ser Glu Asn Ile Arg Trp Ala Glu Ser Val Glu Asn Phe Arg Ser
65 70 75
Gln Gly Val Thr Leu Cys Gly Asp Val Leu Leu Ile Ser Ala Phe
80 85 90
Val Ser Tyr Val Gly Tyr Phe Thr Lys Lys Tyr Arg Asn Glu Leu
95 100 105
Met Glu Lys Phe Trp Ile Pro Tyr Ile His Asn Leu Lys Val Pro
110 115 120
Ile Pro Ile Thr Asn Gly Leu Asp Pro Leu Ser Leu Leu Thr Asp
125 130 135
Asp Ala Asp Val Ala Thr Trp Asn Asn Gln Gly Leu Pro Ser Asp
140 145 150
Arg Met Ser Thr Glu Asn Ala Thr Ile Leu Gly Asn Thr Glu Arg
155 160 165
Trp Pro Leu Ile Val Asp Ala Gln Leu Gln Gly Ile Lys Trp Ile
170 175 180
Lys Asn Lys Tyr Arg Ser Glu Leu Lys Ala Ile Arg Leu Gly Gln
185 190 195
Lys Ser Tyr Leu Asp Val Ile Glu Gln Ala Ile Ser Glu Gly Asp
200 205 210
Thr Leu Leu Ile Glu Asn Ile Gly Glu Thr Val Asp Pro Val Leu
215 220 225
Asp Pro Leu Leu Gly Arg Asn Thr Ile Lys Lys Gly Lys Tyr Ile
230 235 240
Lys Ile Gly Asp Lys Glu Val Glu Tyr His Pro Lys Phe Arg Leu
245 250 255
Ile Leu His Thr Lys Tyr Phe Asn Pro His Tyr Lys Pro Glu Met
260 265 270
Gln Ala Gln Cys Thr Leu Ile Asn Phe Leu Val Thr Arg Asp Gly
275 280 285
Leu Glu Asp Gln Leu Leu Ala Ala Val Val Ala Lys Glu Arg Pro
290 295 300
Asp Leu Glu Gln Leu Lys Ala Asn Leu Thr Lys Ser Gln Asn Glu
305 310 315
Phe Lys Ile Val Leu Lys Glu Leu Glu Asp Ser Leu Leu Ala Arg
320 325 330
Leu Ser Ala Ala Ser Gly Asn Phe Leu Gly Asp Thr Ala Leu Val
335 340 345
Glu Asn Leu Glu Thr Thr Lys His Thr Ala Ser Glu Ile Glu Glu
350 355 360
Lys Val Val Glu Ala Lys Ile Thr Glu Val Lys Ile Asn Glu Ala
365 370 375
Arg Glu Asn Tyr Arg Pro Ala Ala Glu Arg Ala Ser Leu Leu Tyr
380 385 390
Phe Ile Leu Asn Asp Leu Asn Lys Ile Asn Pro Val Tyr Gln Phe

| | | | | | |
|-----------------|---------------------|---------------------|-----|--|-----|
| | 395 | | 400 | | 405 |
| Ser Leu Lys Ala | Phe Asn Val Val Phe | Glu Lys Ala Ile Gln | Arg | | |
| | 410 | | 415 | | 420 |
| Thr Thr Pro Ala | Asn Glu Val Lys Gln | Arg Val Ile Asn Leu | Thr | | |
| | 425 | | 430 | | 435 |
| Asp Glu Ile Thr | Tyr Ser Val Tyr Met | Tyr Thr Ala Arg Gly | Leu | | |
| | 440 | | 445 | | 450 |
| Phe Glu Arg Asp | Lys Leu Ile Phe Leu | Ala Gln Val Thr Phe | Gln | | |
| | 455 | | 460 | | 465 |
| Val Leu Ser Met | Lys Lys Glu Leu Asn | Pro Val Glu Leu Asp | Phe | | |
| | 470 | | 475 | | 480 |
| Leu Leu Arg Phe | Pro Phe Lys Ala Gly | Val Val Ser Pro Val | Asp | | |
| | 485 | | 490 | | 495 |
| Phe Leu Gln His | Gln Gly Trp Gly Gly | Ile Lys Ala Leu Ser | Glu | | |
| | 500 | | 505 | | 510 |
| Met Asp Glu Phe | Lys Asn Leu Asp Ser | Asp Ile Glu Gly Ser | Ala | | |
| | 515 | | 520 | | 525 |
| Lys Arg Trp Lys | Lys Leu Val Glu Ser | Glu Ala Pro Glu Lys | Glu | | |
| | 530 | | 535 | | 540 |
| Ile Phe Pro Lys | Glu Trp Lys Asn Lys | Thr Ala Leu Gln Lys | Leu | | |
| | 545 | | 550 | | 555 |
| Cys Met Val Arg | Cys Leu Arg Pro Asp | Arg Met Thr Tyr Ala | Ile | | |
| | 560 | | 565 | | 570 |
| Lys Asn Phe Val | Glu Glu Lys Met Gly | Ser Lys Phe Val Glu | Gly | | |
| | 575 | | 580 | | 585 |
| Arg Ser Val Glu | Phe Ser Lys Ser Tyr | Glu Glu Ser Ser Pro | Ser | | |
| | 590 | | 595 | | 600 |
| Thr Ser Ile Phe | Phe Ile Leu Ser Pro | Gly Val Asp Pro Leu | Lys | | |
| | 605 | | 610 | | 615 |
| Asp Val Glu Ala | Leu Gly Lys Lys Leu | Gly Phe Thr Ile Asp | Asn | | |
| | 620 | | 625 | | 630 |
| Gly Lys Leu His | Asn Val Ser Leu Gly | Gln Gly Gln Glu Val | Val | | |
| | 635 | | 640 | | 645 |
| Ala Glu Asn Ala | Leu Asp Val Ala Ala | Glu Lys Gly His Trp | Val | | |
| | 650 | | 655 | | 660 |
| Ile Leu Gln Asn | Ile His Leu Val Ala | Arg Trp Leu Gly Thr | Leu | | |
| | 665 | | 670 | | 675 |
| Asp Lys Lys Leu | Glu Arg Tyr Ser Thr | Gly Ser His Glu Asp | Tyr | | |
| | 680 | | 685 | | 690 |
| Arg Val Phe Ile | Ser Ala Glu Pro Ala | Pro Ser Pro Glu Thr | His | | |
| | 695 | | 700 | | 705 |
| Ile Ile Pro Gln | Gly Ile Leu Glu Asn | Ala Ile Lys Ile Thr | Asn | | |
| | 710 | | 715 | | 720 |
| Glu Pro Pro Thr | Gly Met Tyr Ala Asn | Leu His Lys Ala Leu | Asp | | |
| | 725 | | 730 | | 735 |
| Leu Phe Thr Gln | Asp Thr Leu Glu Met | Cys Thr Lys Glu Met | Glu | | |
| | 740 | | 745 | | 750 |
| Phe Lys Cys Met | Leu Phe Ala Leu Cys | Tyr Phe His Ala Val | Val | | |
| | 755 | | 760 | | 765 |
| Ala Glu Arg Arg | Lys Phe Gly Ala Gln | Gly Trp Asn Arg Ser | Tyr | | |
| | 770 | | 775 | | 780 |
| Pro Phe Asn Asn | Gly Asp Leu Thr Ile | Ser Ile Asn Val Leu | Tyr | | |
| | 785 | | 790 | | 795 |
| Asn Tyr Leu Glu | Ala Asn Pro Lys Val | Pro Trp Asp Asp Leu | Arg | | |
| | 800 | | 805 | | 810 |
| Tyr Leu Phe Gly | Glu Ile Met Tyr Gly | Gly His Ile Thr Asp | Asp | | |
| | 815 | | 820 | | 825 |
| Trp Asp Arg Arg | Leu Cys Arg Thr Tyr | Leu Ala Glu Tyr Ile | Arg | | |
| | 830 | | 835 | | 840 |
| Thr Glu Met Leu | Glu Gly Asp Val Leu | Leu Ala Pro Gly Phe | Gln | | |
| | 845 | | 850 | | 855 |
| Ile Pro Pro Asn | Leu Asp Tyr Lys Gly | Tyr His Glu Tyr Ile | Asp | | |
| | 860 | | 865 | | 870 |
| Glu Asn Leu Pro | Pro Glu Ser Pro Tyr | Leu Tyr Gly Leu His | Pro | | |
| | 875 | | 880 | | 885 |
| Asn Ala Glu Ile | Gly Phe Leu Thr Val | Thr Ser Glu Lys Leu | Phe | | |
| | 890 | | 895 | | 900 |

```

Arg Thr Val Leu Glu Met Gln Pro Lys Glu Thr Asp Ser Gly Ala
      905      910      915
Gly Thr Gly Val Ser Arg Glu Glu Lys Val Lys Ala Val Leu Asp
      920      925      930
Asp Ile Leu Glu Lys Ile Pro Glu Thr Phe Asn Met Ala Glu Ile
      935      940      945
Met Ala Lys Ala Ala Glu Lys Thr Pro Tyr Val Val Val Ala Phe
      950      955      960
Gln Glu Cys Glu Arg Met Asn Ile Leu Thr Asn Glu Met Arg Arg
      965      970      975
Ser Leu Lys Glu Leu Asn Leu Gly Leu Lys Gly Glu Leu Thr Ile
      980      985      990
Thr Thr Asp Val Glu Asp Leu Ser Thr Ala Leu Phe Tyr Asp Thr
      995      1000      1005
Val Pro Asp Thr Trp Val Ala Arg Ala Tyr Pro Ser Met Met Gly
      1010      1015      1020
Leu Ala Ala Trp Tyr Ala Asp Leu Leu Leu Arg Ile Arg Glu Leu
      1025      1030      1035
Glu Ala Trp Thr Thr Asp Phe Ala Leu Pro Thr Thr Val Trp Leu
      1040      1045      1050
Ala Gly Phe Phe Asn Pro Gln Ser Phe Leu Thr Ala Ile Met Gln
      1055      1060      1065
Ser Met Ala Arg Lys Asn Glu Trp Pro Leu Asp Lys Met Cys Leu
      1070      1075      1080
Ser Val Glu Val Thr Lys Lys Asn Arg Glu Asp Met Thr Ala Pro
      1085      1090      1095
Pro Arg Glu Gly Ser Tyr Val Tyr Gly Leu Phe Met Glu Gly Ala
      1100      1105      1110
Arg Trp Asp Thr Gln Thr Gly Val Ile Ala Glu Ala Arg Leu Lys
      1115      1120      1125
Glu Leu Thr Pro Ala Met Pro Val Ile Phe Ile Lys Ala Ile Pro
      1130      1135      1140
Val Asp Arg Met Glu Thr Lys Asn Ile Tyr Glu Cys Pro Val Tyr
      1145      1150      1155
Lys Thr Arg Ile Arg Gly Pro Thr Tyr Val Trp Thr Phe Asn Leu
      1160      1165      1170
Lys Thr Lys Glu Lys Ala Ala Lys Trp Ile Leu Ala Ala Val Ala
      1175      1180      1185
Leu Leu Leu Gln Val
      1190

```

<210> 7
 <211> 270
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2242106CD1

```

<400> 7
Met Leu Leu Thr Gln Val Val Trp Leu Pro Glu Pro Gly His Ser
  1      5      10      15
His Arg Phe Gln Val Leu Ser Val Ala Thr Asp Gly Lys Val Leu
      20      25      30
Leu Trp Gln Gly Ile Gly Val Gly Gln Leu Gln Leu Thr Glu Gly
      35      40      45
Phe Ala Leu Val Met Gln Gln Leu Pro Arg Ser Thr Lys Leu Lys
      50      55      60
Lys His Pro Arg Gly Glu Thr Glu Val Gly Ala Thr Ala Val Ala
      65      70      75
Phe Ser Ser Phe Asp Pro Arg Leu Phe Ile Leu Gly Thr Glu Gly
      80      85      90
Gly Phe Pro Leu Lys Cys Ser Leu Ala Ala Gly Glu Ala Ala Leu
      95      100      105
Thr Arg Met Pro Ser Ser Val Pro Leu Arg Ala Pro Ala Gln Phe
      110      115      120

```

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Thr | Phe | Ser | Pro | His | Gly | Gly | Pro | Ile | Tyr | Ser | Val | Ser | Cys | Ser | |
| | | | | 125 | | | | | 130 | | | | | 135 | |
| Pro | Phe | His | Arg | Asn | Leu | Phe | Leu | Ser | Ala | Gly | Thr | Asp | Gly | His | |
| | | | | 140 | | | | | 145 | | | | | 150 | |
| Val | His | Leu | Tyr | Ser | Met | Leu | Gln | Ala | Pro | Pro | Leu | Thr | Ser | Leu | |
| | | | | 155 | | | | | 160 | | | | | 165 | |
| Gln | Leu | Ser | Leu | Lys | Tyr | Leu | Phe | Ala | Val | Arg | Trp | Ser | Pro | Val | |
| | | | | 170 | | | | | 175 | | | | | 180 | |
| Arg | Pro | Leu | Val | Phe | Ala | Ala | Ala | Ser | Gly | Lys | Gly | Asp | Val | Gln | |
| | | | | 185 | | | | | 190 | | | | | 195 | |
| Leu | Phe | Asp | Leu | Gln | Lys | Ser | Ser | Gln | Lys | Pro | Thr | Val | Leu | Ile | |
| | | | | 200 | | | | | 205 | | | | | 210 | |
| Lys | Gln | Thr | Gln | Asp | Glu | Ser | Pro | Val | Tyr | Cys | Leu | Glu | Phe | Asn | |
| | | | | 215 | | | | | 220 | | | | | 225 | |
| Ser | Gln | Gln | Thr | Gln | Leu | Leu | Ala | Ala | Gly | Asp | Ala | Gln | Gly | Thr | |
| | | | | 230 | | | | | 235 | | | | | 240 | |
| Val | Lys | Val | Trp | Gln | Leu | Ser | Thr | Glu | Phe | Thr | Glu | Gln | Gly | Pro | |
| | | | | 245 | | | | | 250 | | | | | 255 | |
| Arg | Glu | Ala | Glu | Asp | Leu | Asp | Cys | Leu | Ala | Ala | Glu | Val | Ala | Ala | |
| | | | | 260 | | | | | 265 | | | | | 270 | |

<210> 8
 <211> 647
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2726877CD1

| | | | | | | | | | | | | | | | |
|---------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| <400> 8 | | | | | | | | | | | | | | | |
| Met | Asp | Leu | Leu | Lys | Asn | Pro | Lys | Ile | Ala | Asp | Tyr | Leu | Thr | Arg | |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Tyr | Glu | His | Phe | Ser | Ser | Cys | Leu | His | Gln | Val | Leu | Gly | Leu | Leu | |
| | | | | 20 | | | | | 25 | | | | | 30 | |
| Asn | Gly | Lys | Asp | Pro | Asp | Ser | Ser | Ser | Lys | Val | Leu | Glu | Leu | Leu | |
| | | | | 35 | | | | | 40 | | | | | 45 | |
| Leu | Ala | Phe | Cys | Ser | Val | Thr | Gln | Leu | Arg | His | Met | Leu | Thr | Gln | |
| | | | | 50 | | | | | 55 | | | | | 60 | |
| Met | Met | Phe | Glu | Gln | Ser | Pro | Pro | Gly | Ser | Ala | Thr | Leu | Gly | Ser | |
| | | | | 65 | | | | | 70 | | | | | 75 | |
| His | Thr | Lys | Cys | Leu | Glu | Pro | Thr | Val | Ala | Leu | Leu | Arg | Trp | Leu | |
| | | | | 80 | | | | | 85 | | | | | 90 | |
| Ser | Gln | Pro | Leu | Asp | Gly | Ser | Glu | Asn | Cys | Ser | Val | Leu | Ala | Leu | |
| | | | | 95 | | | | | 100 | | | | | 105 | |
| Glu | Leu | Phe | Lys | Glu | Ile | Phe | Glu | Asp | Val | Ile | Asp | Ala | Ala | Asn | |
| | | | | 110 | | | | | 115 | | | | | 120 | |
| Cys | Ser | Ser | Ala | Asp | Arg | Phe | Val | Thr | Leu | Leu | Leu | Pro | Thr | Ile | |
| | | | | 125 | | | | | 130 | | | | | 135 | |
| Leu | Asp | Gln | Leu | Gln | Phe | Thr | Glu | Gln | Asn | Leu | Asp | Glu | Ala | Leu | |
| | | | | 140 | | | | | 145 | | | | | 150 | |
| Thr | Arg | Gln | Lys | Cys | Glu | Arg | Ile | Ala | Lys | Ala | Phe | Glu | Val | Leu | |
| | | | | 155 | | | | | 160 | | | | | 165 | |
| Leu | Thr | Leu | Cys | Gly | Asp | Asp | Thr | Leu | Lys | Met | His | Ile | Ala | Lys | |
| | | | | 170 | | | | | 175 | | | | | 180 | |
| Ile | Leu | Thr | Thr | Val | Lys | Cys | Thr | Thr | Leu | Ile | Glu | Gln | Gln | Phe | |
| | | | | 185 | | | | | 190 | | | | | 195 | |
| Thr | Tyr | Gly | Lys | Ile | Asp | Leu | Gly | Phe | Gly | Thr | Lys | Val | Ala | Asp | |
| | | | | 200 | | | | | 205 | | | | | 210 | |
| Ser | Glu | Leu | Cys | Lys | Leu | Ala | Ala | Asp | Val | Ile | Leu | Lys | Thr | Leu | |
| | | | | 215 | | | | | 220 | | | | | 225 | |
| Asp | Leu | Ile | Asn | Lys | Leu | Lys | Pro | Leu | Val | Pro | Gly | Met | Glu | Val | |
| | | | | 230 | | | | | 235 | | | | | 240 | |
| Ser | Phe | Tyr | Lys | Ile | Leu | Gln | Asp | Pro | Arg | Leu | Ile | Thr | Pro | Leu | |
| | | | | 245 | | | | | 250 | | | | | 255 | |
| Ala | Phe | Ala | Leu | Thr | Ser | Asp | Asn | Arg | Glu | Gln | Val | Gln | Ser | Gly | |

| | | | | | |
|-----------------|-----|---------------------|-----|---------------------|-----|
| Leu Arg Ile Leu | 260 | Leu Glu Ala Ala Pro | 265 | Leu Pro Asp Phe Pro | 270 |
| | 275 | | 280 | | 285 |
| Leu Val Leu Gly | 290 | Glu Ser Ile Ala Ala | 295 | Asn Asn Ala Tyr Arg | 300 |
| Gln Glu Thr Glu | 305 | His Ile Pro Arg Lys | 310 | Met Pro Trp Gln Ser | 315 |
| Asn His Ser Phe | 320 | Pro Thr Ser Ile Lys | 325 | Cys Leu Thr Pro His | 330 |
| Lys Asp Gly Val | 335 | Pro Gly Leu Asn Ile | 340 | Glu Glu Leu Ile Glu | 345 |
| Leu Gln Ser Gly | 350 | Met Val Val Lys Asp | 355 | Gln Ile Cys Asp Val | 360 |
| Ile Ser Asp Ile | 365 | Met Asp Val Tyr Glu | 370 | Met Lys Leu Ser Thr | 375 |
| Ala Ser Lys Glu | 380 | Ser Arg Leu Gln Asp | 385 | Leu Leu Glu Thr Lys | 390 |
| Leu Ala Leu Ala | 395 | Gln Ala Asp Arg Leu | 400 | Ile Ala Gln His Arg | 405 |
| Gln Arg Thr Gln | 410 | Ala Glu Thr Glu Ala | 415 | Arg Thr Leu Ala Ser | 420 |
| Leu Arg Glu Val | 425 | Glu Arg Lys Asn Glu | 430 | Glu Leu Ser Val Leu | 435 |
| Lys Ala Gln Gln | 440 | Val Glu Ser Glu Arg | 445 | Ala Gln Ser Asp Ile | 450 |
| His Leu Phe Gln | 455 | His Asn Arg Lys Leu | 460 | Glu Ser Val Ala Glu | 465 |
| His Glu Ile Leu | 470 | Thr Lys Ser Tyr Met | 475 | Glu Leu Leu Gln Arg | 480 |
| Glu Ser Thr Glu | 485 | Lys Lys Asn Lys Asp | 490 | Leu Gln Ile Thr Cys | 495 |
| Ser Leu Asn Lys | 500 | Gln Ile Glu Thr Val | 505 | Lys Lys Leu Asn Glu | 510 |
| Leu Lys Glu Gln | 515 | Asn Glu Lys Ser Ile | 520 | Ala Gln Leu Ile Glu | 525 |
| Glu Glu Gln Arg | 530 | Lys Glu Val Gln Asn | 535 | Gln Leu Val Asp Arg | 540 |
| His Lys Leu Ala | 545 | Asn Leu His Gln Lys | 550 | Thr Lys Val Gln Glu | 555 |
| Lys Ile Lys Thr | 560 | Leu Gln Lys Glu Arg | 565 | Glu Asp Lys Glu Glu | 570 |
| Ile Asp Ile Leu | 575 | Arg Lys Glu Leu Ser | 580 | Arg Thr Glu Gln Ile | 585 |
| Lys Glu Leu Ser | 590 | Ile Lys Ala Ser Ser | 595 | Leu Glu Val Gln Lys | 600 |
| Gln Leu Glu Gly | 605 | Arg Leu Glu Glu Lys | 610 | Glu Ser Leu Val Lys | 615 |
| Gln Gln Glu Glu | 620 | Leu Asn Lys His Ser | 625 | His Met Ile Ala Met | 630 |
| His Ser Leu Ser | 635 | Gly Lys Ile Asn | 640 | Pro Glu Thr Val Asn | 645 |
| Ser Ile | | | | | |

<210> 9

<211> 1086

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2738233CD1

<400> 9

| | | |
|-----------------|-------------------------|---------------------|
| Met Arg Cys Lys | Glu Leu Glu Asn Ala Val | Gly Ser Trp Thr Asp |
| 1 | 5 | 10 |
| Asp Leu Thr Gln | Leu Ser Leu Leu Lys Asp | Thr Leu Ser Ala Tyr |

| | | | | | |
|---------------------|-----|---------------------|-----|---------------------|-----|
| Ile Ser Ala Asp | 20 | Ile Ser Ile Leu | 25 | Glu Arg Val Glu | 30 |
| Leu Gln Arg Gln Trp | 35 | Glu Glu Leu Cys His | 40 | Gln Leu Ser Leu Arg | 45 |
| Arg Gln Gln Ile Gly | 50 | Glu Arg Leu Asn Glu | 55 | Trp Ala Val Phe Ser | 60 |
| Glu Lys Asn Lys Glu | 65 | Leu Cys Glu Trp Leu | 70 | Thr Gln Met Glu Ser | 75 |
| Lys Val Ser Gln Asn | 80 | Gly Asp Ile Leu Ile | 85 | Glu Glu Met Ile Glu | 90 |
| Lys Leu Lys Lys Asp | 95 | Tyr Gln Glu Glu Ile | 100 | Ala Ile Ala Gln Glu | 105 |
| Asn Lys Ile Gln Leu | 110 | Gln Gln Met Gly Glu | 115 | Arg Leu Ala Lys Ala | 120 |
| Ser His Glu Ser Lys | 125 | Ala Ser Glu Ile Glu | 130 | Tyr Lys Leu Gly Lys | 135 |
| Val Asn Asp Arg Trp | 140 | Gln His Leu Leu Asp | 145 | Leu Ile Ala Ala Arg | 150 |
| Val Lys Lys Leu Lys | 155 | Glu Thr Leu Val Ala | 160 | Val Gln Gln Leu Asp | 165 |
| Lys Asn Met Ser Ser | 170 | Leu Arg Thr Trp Leu | 175 | Ala His Ile Glu Ser | 180 |
| Glu Leu Ala Lys Pro | 185 | Ile Val Tyr Gly Ser | 190 | Cys Asn Ser Glu Glu | 195 |
| Ile Gln Arg Lys Leu | 200 | Asn Glu Gln Gln Glu | 205 | Leu Gln Arg Asp Ile | 210 |
| Glu Lys His Ser Thr | 215 | Gly Val Ala Ser Val | 220 | Leu Asn Leu Cys Glu | 225 |
| Val Leu Leu His Asp | 230 | Cys Asp Ala Cys Ala | 235 | Thr Asp Ala Glu Cys | 240 |
| Asp Ser Ile Gln Gln | 245 | Ala Thr Arg Asn Leu | 250 | Asp Arg Arg Trp Arg | 255 |
| Asn Ile Cys Ala Met | 260 | Ser Met Glu Arg Arg | 265 | Leu Lys Ile Glu Glu | 270 |
| Thr Trp Arg Leu Trp | 275 | Gln Lys Phe Leu Asp | 280 | Asp Tyr Ser Arg Phe | 285 |
| Glu Asp Trp Leu Lys | 290 | Ser Ser Glu Arg Thr | 295 | Ala Ala Phe Pro Ser | 300 |
| Ser Ser Gly Val Ile | 305 | Tyr Thr Val Ala Lys | 310 | Glu Glu Leu Lys Lys | 315 |
| Phe Glu Ala Phe Gln | 320 | Arg Gln Val His Glu | 325 | Cys Leu Thr Gln Leu | 330 |
| Glu Leu Ile Asn Lys | 335 | Gln Tyr Arg Arg Leu | 340 | Ala Arg Glu Asn Arg | 345 |
| Thr Asp Ser Ala Cys | 350 | Ser Leu Lys Gln Met | 355 | Val His Glu Gly Asn | 360 |
| Gln Arg Trp Asp Asn | 365 | Leu Gln Lys Arg Val | 370 | Thr Ser Ile Leu Arg | 375 |
| Arg Leu Lys His Phe | 380 | Ile Gly Gln Arg Glu | 385 | Glu Phe Glu Thr Ala | 390 |
| Arg Asp Ser Ile Leu | 395 | Val Trp Leu Thr Glu | 400 | Met Asp Leu Gln Leu | 405 |
| Thr Asn Ile Glu His | 410 | Phe Ser Glu Cys Asp | 415 | Val Gln Ala Lys Ile | 420 |
| Lys Gln Leu Lys Ala | 425 | Phe Gln Gln Glu Ile | 430 | Ser Leu Asn His Asn | 435 |
| Lys Ile Glu Gln Ile | 440 | Ile Ala Gln Gly Glu | 445 | Gln Leu Ile Glu Lys | 450 |
| Ser Glu Pro Leu Asp | 455 | Ala Ala Ile Ile Glu | 460 | Glu Glu Leu Asp Glu | 465 |
| Leu Arg Arg Tyr Cys | 470 | Gln Glu Val Phe Gly | 475 | Arg Val Glu Arg Tyr | 480 |
| His Lys Lys Leu Ile | 485 | Arg Leu Pro Leu Pro | 490 | Asp Asp Glu His Asp | 495 |
| Leu Ser Asp Arg Glu | 500 | Leu Glu Leu Glu Asp | 505 | Ser Ala Ala Leu Ser | 510 |
| | 515 | | 520 | | 525 |

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|------|-----|-----|-----|-----|------|-----|-----|-----|-----|------|
| Asp | Leu | His | Trp | His | Asp | Arg | Ser | Ala | Asp | Ser | Leu | Leu | Ser | Pro |
| | | | | 530 | | | | | 535 | | | | | 540 |
| Gln | Pro | Ser | Ser | Asn | Leu | Ser | Leu | Ser | Leu | Ala | Gln | Pro | Leu | Arg |
| | | | | 545 | | | | | 550 | | | | | 555 |
| Ser | Glu | Arg | Ser | Gly | Arg | Asp | Thr | Pro | Ala | Ser | Val | Asp | Ser | Ile |
| | | | | 560 | | | | | 565 | | | | | 570 |
| Pro | Leu | Glu | Trp | Asp | His | Asp | Tyr | Asp | Leu | Ser | Arg | Asp | Leu | Glu |
| | | | | 575 | | | | | 580 | | | | | 585 |
| Ser | Ala | Met | Ser | Arg | Ala | Leu | Pro | Ser | Glu | Asp | Glu | Glu | Gly | Gln |
| | | | | 590 | | | | | 595 | | | | | 600 |
| Asp | Asp | Lys | Asp | Phe | Tyr | Leu | Arg | Gly | Ala | Val | Gly | Leu | Ser | Gly |
| | | | | 605 | | | | | 610 | | | | | 615 |
| Asp | His | Ser | Ala | Leu | Glu | Ser | Gln | Ile | Arg | Gln | Leu | Gly | Lys | Ala |
| | | | | 620 | | | | | 625 | | | | | 630 |
| Leu | Asp | Asp | Ser | Arg | Phe | Gln | Ile | Gln | Gln | Thr | Glu | Asn | Ile | Ile |
| | | | | 635 | | | | | 640 | | | | | 645 |
| Arg | Ser | Lys | Thr | Pro | Thr | Gly | Pro | Glu | Leu | Asp | Thr | Ser | Tyr | Lys |
| | | | | 650 | | | | | 655 | | | | | 660 |
| Gly | Tyr | Met | Lys | Leu | Leu | Gly | Glu | Cys | Ser | Ser | Ser | Ile | Asp | Ser |
| | | | | 665 | | | | | 670 | | | | | 675 |
| Val | Lys | Arg | Leu | Glu | His | Lys | Leu | Lys | Glu | Glu | Glu | Glu | Ser | Leu |
| | | | | 680 | | | | | 685 | | | | | 690 |
| Pro | Gly | Phe | Val | Asn | Leu | His | Ser | Thr | Glu | Thr | Gln | Thr | Ala | Gly |
| | | | | 695 | | | | | 700 | | | | | 705 |
| Val | Ile | Asp | Arg | Trp | Glu | Leu | Leu | Gln | Ala | Gln | Ala | Leu | Ser | Lys |
| | | | | 710 | | | | | 715 | | | | | 720 |
| Glu | Leu | Arg | Met | Lys | Gln | Asn | Leu | Gln | Lys | Trp | Gln | Gln | Phe | Asn |
| | | | | 725 | | | | | 730 | | | | | 735 |
| Ser | Asp | Leu | Asn | Ser | Ile | Trp | Ala | Trp | Leu | Gly | Asp | Thr | Glu | Glu |
| | | | | 740 | | | | | 745 | | | | | 750 |
| Glu | Leu | Glu | Gln | Leu | Gln | Arg | Leu | Glu | Leu | Ser | Thr | Asp | Ile | Gln |
| | | | | 755 | | | | | 760 | | | | | 765 |
| Thr | Ile | Glu | Leu | Gln | Ile | Lys | Lys | Leu | Lys | Glu | Leu | Gln | Lys | Ala |
| | | | | 770 | | | | | 775 | | | | | 780 |
| Val | Asp | His | Arg | Lys | Ala | Ile | Ile | Leu | Ser | Ile | Asn | Leu | Cys | Ser |
| | | | | 785 | | | | | 790 | | | | | 795 |
| Pro | Glu | Phe | Thr | Gln | Ala | Asp | Ser | Lys | Glu | Ser | Arg | Asp | Leu | Gln |
| | | | | 800 | | | | | 805 | | | | | 810 |
| Asp | Arg | Leu | Ser | Gln | Met | Asn | Gly | Arg | Trp | Asp | Arg | Val | Cys | Ser |
| | | | | 815 | | | | | 820 | | | | | 825 |
| Leu | Leu | Glu | Glu | Trp | Arg | Gly | Leu | Leu | Gln | Asp | Ala | Leu | Met | Gln |
| | | | | 830 | | | | | 835 | | | | | 840 |
| Cys | Gln | Gly | Phe | His | Glu | Met | Ser | His | Gly | Leu | Leu | Leu | Met | Leu |
| | | | | 845 | | | | | 850 | | | | | 855 |
| Glu | Asn | Ile | Asp | Arg | Arg | Lys | Asn | Glu | Ile | Val | Pro | Ile | Asp | Ser |
| | | | | 860 | | | | | 865 | | | | | 870 |
| Asn | Leu | Asp | Ala | Glu | Ile | Leu | Gln | Asp | His | His | Lys | Gln | Leu | Met |
| | | | | 875 | | | | | 880 | | | | | 885 |
| Gln | Ile | Lys | His | Glu | Leu | Leu | Glu | Ser | Gln | Leu | Arg | Val | Ala | Ser |
| | | | | 890 | | | | | 895 | | | | | 900 |
| Leu | Gln | Asp | Met | Ser | Cys | Gln | Leu | Leu | Val | Asn | Ala | Glu | Gly | Thr |
| | | | | 905 | | | | | 910 | | | | | 915 |
| Asp | Cys | Leu | Glu | Ala | Lys | Glu | Lys | Val | His | Val | Ile | Gly | Asn | Arg |
| | | | | 920 | | | | | 925 | | | | | 930 |
| Leu | Lys | Leu | Leu | Leu | Lys | Glu | Val | Ser | Arg | His | Ile | Lys | Glu | Leu |
| | | | | 935 | | | | | 940 | | | | | 945 |
| Glu | Lys | Leu | Leu | Asp | Val | Ser | Ser | Ser | Gln | Gln | Asp | Leu | Ser | Ser |
| | | | | 950 | | | | | 955 | | | | | 960 |
| Trp | Ser | Ser | Ala | Asp | Glu | Leu | Asp | Thr | Ser | Gly | Ser | Val | Ser | Pro |
| | | | | 965 | | | | | 970 | | | | | 975 |
| Thr | Ser | Gly | Arg | Ser | Thr | Pro | Asn | Arg | Gln | Lys | Thr | Pro | Arg | Gly |
| | | | | 980 | | | | | 985 | | | | | 990 |
| Lys | Cys | Ser | Leu | Ser | Gln | Pro | Gly | Pro | Ser | Val | Ser | Ser | Pro | His |
| | | | | 995 | | | | | 1000 | | | | | 1005 |
| Ser | Arg | Ser | Thr | Lys | Gly | Gly | Ser | Asp | Ser | Ser | Leu | Ser | Glu | Pro |
| | | | | 1010 | | | | | 1015 | | | | | 1020 |
| Gly | Pro | Gly | Arg | Ser | Gly | Arg | Gly | Phe | Leu | Phe | Arg | Val | Leu | Arg |

| | | |
|---------------------|---|------|
| 1025 | 1030 | 1035 |
| Ala Ala Leu Pro Leu | Gln Leu Leu Leu Leu Leu Ile Gly Leu | |
| 1040 | 1045 | 1050 |
| Ala Cys Leu Val Pro | Met Ser Glu Glu Asp Tyr Ser Cys Ala Leu | |
| 1055 | 1060 | 1065 |
| Ser Asn Asn Phe Ala | Arg Ser Phe His Pro Met Leu Arg Tyr Thr | |
| 1070 | 1075 | 1080 |
| Asn Gly Pro Pro Pro | Leu | |
| 1085 | | |

<210> 10

<211> 396

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1833116CD1

<400> 10

| | | |
|---------------------|---|-----|
| Met Thr Thr Leu Val | Leu Asp Asn Gly Ala Tyr Asn Ala Lys Ile | |
| 1 | 5 | 10 |
| Gly Tyr Ser His Glu | Asn Val Ser Val Ile Pro Asn Cys Gln Phe | 15 |
| 20 | 25 | 30 |
| Arg Ser Lys Thr Ala | Arg Leu Lys Thr Phe Thr Ala Asn Gln Ile | 35 |
| 35 | 40 | 45 |
| Asp Glu Ile Lys Asp | Pro Ser Gly Leu Phe Tyr Ile Leu Pro Phe | 50 |
| 50 | 55 | 60 |
| Gln Lys Gly Tyr Leu | Val Asn Trp Asp Val Gln Arg Gln Val Trp | 65 |
| 65 | 70 | 75 |
| Asp Tyr Leu Phe Gly | Lys Glu Met Tyr Gln Val Asp Phe Leu Asp | 80 |
| 80 | 85 | 90 |
| Thr Asn Ile Ile Ile | Thr Glu Pro Tyr Phe Asn Phe Thr Ser Ile | 95 |
| 95 | 100 | 105 |
| Gln Glu Ser Met Asn | Glu Ile Leu Phe Glu Glu Tyr Gln Phe Gln | 110 |
| 110 | 115 | 120 |
| Ala Val Leu Arg Val | Asn Ala Gly Ala Leu Ser Ala His Arg Tyr | 125 |
| 125 | 130 | 135 |
| Phe Arg Asp Asn Pro | Ser Glu Leu Cys Cys Ile Ile Val Asp Ser | 140 |
| 140 | 145 | 150 |
| Gly Tyr Ser Phe Thr | His Ile Val Pro Tyr Cys Arg Ser Lys Lys | 155 |
| 155 | 160 | 165 |
| Lys Lys Glu Ala Ile | Ile Arg Ile Asn Val Gly Gly Lys Leu Leu | 170 |
| 170 | 175 | 180 |
| Thr Asn His Leu Lys | Glu Ile Ile Ser Tyr Arg Gln Leu His Val | 185 |
| 185 | 190 | 195 |
| Met Asp Glu Thr His | Val Ile Asn Gln Val Lys Glu Asp Val Cys | 200 |
| 200 | 205 | 210 |
| Tyr Val Ser Gln Asp | Phe Tyr Arg Asp Met Asp Ile Ala Lys Leu | 215 |
| 215 | 220 | 225 |
| Lys Gly Glu Glu Asn | Thr Val Met Ile Asp Tyr Val Leu Pro Asp | 230 |
| 230 | 235 | 240 |
| Phe Ser Thr Ile Lys | Lys Gly Phe Cys Lys Pro Arg Glu Glu Met | 245 |
| 245 | 250 | 255 |
| Val Leu Ser Gly Lys | Tyr Lys Ser Gly Glu Gln Ile Leu Arg Leu | 260 |
| 260 | 265 | 270 |
| Ala Asn Glu Arg Phe | Ala Val Pro Glu Ile Leu Phe Asn Pro Ser | 275 |
| 275 | 280 | 285 |
| Asp Ile Gly Ile Gln | Glu Met Gly Ile Pro Glu Ala Ile Val Tyr | 290 |
| 290 | 295 | 300 |
| Ser Ile Gln Asn Leu | Pro Glu Glu Met Gln Pro His Phe Phe Lys | 305 |
| 305 | 310 | 315 |
| Asn Ile Val Leu Thr | Gly Gly Asn Ser Leu Phe Pro Gly Phe Arg | 320 |
| 320 | 325 | 330 |
| Asp Arg Val Tyr Ser | Glu Val Arg Cys Leu Thr Pro Thr Asp Tyr | 335 |
| 335 | 340 | 345 |
| Asp Val Ser Val Val | Leu Pro Glu Asn Pro Ile Thr Tyr Ala Trp | |

| | | | |
|---|-----|-----|-----|
| Glu Gly Gly Lys | 350 | 355 | 360 |
| Leu Ile Ser Glu Asn Asp Asp Phe Glu Asp Met | 365 | 370 | 375 |
| Val Val Thr Arg Glu Asp Tyr Glu Glu Asn Gly His Ser Val Cys | 380 | 385 | 390 |
| Glu Glu Lys Phe Asp Ile | 395 | | |

<210> 11
 <211> 304
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 001799CD1

<400> 11

| | | |
|---------------------|---|-----|
| Met Ala Ser Glu Thr | His Asn Val Lys Lys Arg Asn Phe Cys Asn | |
| 1 | 5 | 10 |
| Lys Ile Glu Asp His | Phe Ile Asp Leu Pro Arg Lys Lys Ile Ser | 15 |
| | 20 | 25 |
| Asn Phe Thr Asn Lys | Asn Met Lys Glu Val Lys Lys Ser Pro Lys | 30 |
| | 35 | 40 |
| Gln Leu Ala Ala Tyr | Ile Asn Arg Thr Val Gly Gln Thr Val Lys | 45 |
| | 50 | 55 |
| Ser Pro Asp Lys Leu | Arg Lys Val Ile Tyr Arg Arg Lys Lys Val | 60 |
| | 65 | 70 |
| His His Pro Phe Pro | Asn Pro Cys Tyr Arg Lys Lys Gln Ser Pro | 75 |
| | 80 | 85 |
| Gly Ser Gly Gly Cys | Asp Met Ala Asn Lys Glu Asn Glu Leu Ala | 90 |
| | 95 | 100 |
| Cys Ala Gly His Leu | Pro Glu Lys Leu His His Asp Ser Arg Thr | 105 |
| | 110 | 115 |
| Tyr Leu Val Asn Ser | Ser Asp Ser Gly Ser Ser Gln Thr Glu Ser | 120 |
| | 125 | 130 |
| Pro Ser Ser Lys Tyr | Ser Gly Phe Phe Ser Glu Val Ser Gln Asp | 135 |
| | 140 | 145 |
| His Glu Thr Met Ala | Gln Val Leu Phe Ser Arg Asn Met Arg Leu | 150 |
| | 155 | 160 |
| Asn Val Ala Leu Thr | Phe Trp Arg Lys Arg Ser Ile Ser Glu Leu | 165 |
| | 170 | 175 |
| Val Ala Tyr Leu Leu | Arg Ile Glu Asp Leu Gly Val Val Val Asp | 180 |
| | 185 | 190 |
| Cys Leu Pro Val Leu | Thr Asn Cys Leu Gln Glu Glu Lys Gln Tyr | 195 |
| | 200 | 205 |
| Ile Ser Leu Gly Cys | Cys Val Asp Leu Leu Pro Leu Val Lys Ser | 210 |
| | 215 | 220 |
| Leu Leu Lys Ser Lys | Phe Glu Glu Tyr Val Ile Val Gly Leu Asn | 225 |
| | 230 | 235 |
| Trp Leu Gln Ala Val | Ile Lys Arg Trp Trp Ser Glu Leu Ser Ser | 240 |
| | 245 | 250 |
| Lys Thr Glu Ile Ile | Asn Asp Gly Asn Ile Gln Ile Leu Lys Gln | 255 |
| | 260 | 265 |
| Gln Leu Ser Gly Leu | Trp Glu Gln Glu Asn His Leu Thr Leu Val | 270 |
| | 275 | 280 |
| Pro Gly Tyr Thr Gly | Asn Ile Ala Lys Asp Val Asp Ala Tyr Leu | 285 |
| | 290 | 295 |
| Leu Gln Leu His | | 300 |

<210> 12
 <211> 201
 <212> PRT
 <213> Homo sapiens

<220>

<400> 12

```
<210> 13
<211> 547
<212> PRT
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<223> Incyte ID No: 1295420CD1
```

<400> 13

15/69

| | | | | | | | | | | | | | | |
|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| <400> 14 | | | | | | | | | | | | | | |
| Met | Glu | Thr | Leu | Ser | Phe | Pro | Arg | Tyr | Asn | Val | Ala | Glu | Ile | Val |
| 1 | | | | 5 | | | | | 10 | | | | | 15 |
| Ile | His | Ile | Arg | Asn | Lys | Ile | Leu | Thr | Gly | Ala | Asp | Gly | Lys | Asn |
| | | | | 20 | | | | | 25 | | | | | 30 |
| Leu | Thr | Lys | Asn | Asp | Leu | Tyr | Pro | Asn | Pro | Lys | Pro | Glu | Val | Leu |
| | | | | 35 | | | | | 40 | | | | | 45 |

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| His | Met | Ile | Tyr | Met | Arg | Ala | Leu | Gln | Ile | Val | Tyr | Gly | Ile | Arg |
| | | | | 50 | | | | | 55 | | | | | 60 |
| Leu | Glu | His | Phe | Tyr | Met | Met | Pro | Val | Asn | Ser | Glu | Val | Met | Tyr |
| | | | | 65 | | | | | 70 | | | | | 75 |
| Pro | His | Leu | Met | Glu | Gly | Phe | Leu | Pro | Phe | Ser | Asn | Leu | Val | Thr |
| | | | | 80 | | | | | 85 | | | | | 90 |
| His | Leu | Asp | Ser | Phe | Leu | Pro | Ile | Cys | Arg | Val | Asn | Asp | Phe | Glu |
| | | | | 95 | | | | | 100 | | | | | 105 |
| Thr | Ala | Asp | Ile | Leu | Cys | Pro | Lys | Ala | Lys | Arg | Thr | Ser | Arg | Phe |
| | | | | 110 | | | | | 115 | | | | | 120 |
| Leu | Ser | Gly | Ile | Ile | Asn | Phe | Ile | His | Phe | Arg | Glu | Ala | Cys | Arg |
| | | | | 125 | | | | | 130 | | | | | 135 |
| Glu | Thr | Tyr | Met | Glu | Phe | Leu | Trp | Gln | Tyr | Lys | Ser | Ser | Ala | Asp |
| | | | | 140 | | | | | 145 | | | | | 150 |
| Lys | Met | Gln | Gln | Leu | Asn | Ala | Ala | His | Gln | Glu | Ala | Leu | Met | Lys |
| | | | | 155 | | | | | 160 | | | | | 165 |
| Leu | Glu | Arg | Leu | Asp | Ser | Val | Pro | Val | Glu | Glu | Gln | Glu | Glu | Phe |
| | | | | 170 | | | | | 175 | | | | | 180 |
| Lys | Gln | Leu | Ser | Asp | Gly | Ile | Gln | Glu | Leu | Gln | Gln | Ser | Leu | Asn |
| | | | | 185 | | | | | 190 | | | | | 195 |
| Gln | Asp | Phe | His | Gln | Lys | Thr | Ile | Val | Leu | Gln | Glu | Gly | Asn | Pro |
| | | | | 200 | | | | | 205 | | | | | 210 |
| Gln | Lys | Lys | Ser | Asn | Ile | Ser | Glu | Lys | Thr | Lys | Arg | Leu | Asn | Glu |
| | | | | 215 | | | | | 220 | | | | | 225 |
| Leu | Lys | Leu | Leu | Val | Val | Ser | Leu | Lys | Glu | Ile | Gln | Glu | Ser | Leu |
| | | | | 230 | | | | | 235 | | | | | 240 |
| Lys | Thr | Lys | Ile | Val | Asp | Ser | Pro | Glu | Lys | Leu | Lys | Asn | Tyr | Lys |
| | | | | 245 | | | | | 250 | | | | | 255 |
| Glu | Lys | Met | Lys | Asp | Thr | Val | Gln | Lys | Leu | Lys | Asn | Ala | Arg | Gln |
| | | | | 260 | | | | | 265 | | | | | 270 |
| Glu | Val | Val | Glu | Lys | Tyr | Glu | Ile | Tyr | Gly | Asp | Ser | Val | Asp | Cys |
| | | | | 275 | | | | | 280 | | | | | 285 |
| Leu | Pro | Ser | Cys | Gln | Leu | Glu | Val | Gln | Leu | Tyr | Gln | Lys | Lys | Ile |
| | | | | 290 | | | | | 295 | | | | | 300 |
| Gln | Asp | Leu | Ser | Asp | Asn | Arg | Glu | Lys | Leu | Ala | Ser | Ile | Leu | Lys |
| | | | | 305 | | | | | 310 | | | | | 315 |
| Glu | Ser | Leu | Asn | Leu | Glu | Asp | Gln | Ile | Glu | Ser | Asp | Glu | Ser | Glu |
| | | | | 320 | | | | | 325 | | | | | 330 |
| Leu | Lys | Lys | Leu | Lys | Thr | Glu | Glu | Asn | Ser | Phe | Lys | Arg | Leu | Met |
| | | | | 335 | | | | | 340 | | | | | 345 |
| Ile | Val | Lys | Lys | Glu | Lys | Leu | Ala | Thr | Ala | Gln | Phe | Lys | Ile | Asn |
| | | | | 350 | | | | | 355 | | | | | 360 |
| Lys | Lys | His | Glu | Asp | Val | Lys | Gln | Tyr | Lys | Arg | Thr | Val | Ile | Glu |
| | | | | 365 | | | | | 370 | | | | | 375 |
| Asp | Cys | Asn | Lys | Val | Gln | Glu | Lys | Arg | Gly | Ala | Val | Tyr | Glu | Arg |
| | | | | 380 | | | | | 385 | | | | | 390 |
| Val | Thr | Thr | Ile | Asn | Gln | Glu | Ile | Gln | Lys | Ile | Lys | Leu | Gly | Ile |
| | | | | 395 | | | | | 400 | | | | | 405 |
| Gln | Gln | Leu | Lys | Asp | Ala | Ala | Glu | Arg | Glu | Lys | Leu | Lys | Ser | Gln |
| | | | | 410 | | | | | 415 | | | | | 420 |
| Glu | Ile | Phe | Leu | Asn | Leu | Lys | Thr | Ala | Leu | Glu | Lys | Tyr | His | Asp |
| | | | | 425 | | | | | 430 | | | | | 435 |
| Gly | Ile | Glu | Lys | Ala | Ala | Glu | Asp | Ser | Tyr | Ala | Lys | Ile | Asp | Glu |
| | | | | 440 | | | | | 445 | | | | | 450 |
| Lys | Thr | Ala | Glu | Leu | Lys | Arg | Lys | Met | Phe | Lys | Met | Ser | Thr | |
| | | | | 455 | | | | | 460 | | | | | |

<210> 15

<211> 569

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1315267CD1

<400> 15

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Glu | Ser | Ile | Lys | His | Lys | Val | Ser | Glu | Pro | Ser | Arg | Ser | Ser | 15 |
| 1 | | | | 5 | | | | | 10 | | | | | | |
| Ser | Leu | Ser | Leu | Ser | Lys | Met | Asp | Phe | Asp | Asp | Glu | Arg | Thr | Trp | 20 |
| | | | | 20 | | | | | 25 | | | | | | 30 |
| Thr | Asp | Leu | Glu | Glu | Asn | Leu | Cys | Asn | His | Asp | Val | Val | Leu | Gly | 35 |
| | | | | 35 | | | | | 40 | | | | | | 45 |
| Asn | Glu | Ser | Thr | Tyr | Gly | Thr | Pro | Gln | Thr | Cys | Tyr | Pro | Asn | Asn | 50 |
| | | | | 50 | | | | | 55 | | | | | | 60 |
| Glu | Ile | Gly | Ile | Leu | Asp | Lys | Thr | Ile | Lys | Arg | Lys | Ile | Ala | Pro | 65 |
| | | | | 65 | | | | | 70 | | | | | | 75 |
| Val | Lys | Arg | Gly | Glu | Asp | Leu | Ser | Lys | Ser | Arg | Arg | Ser | Arg | Ser | 80 |
| | | | | 80 | | | | | 85 | | | | | | 90 |
| Pro | Pro | Thr | Ser | Glu | Leu | Met | Met | Lys | Phe | Phe | Pro | Ser | Leu | Lys | 95 |
| | | | | 95 | | | | | 100 | | | | | | 105 |
| Pro | Lys | Pro | Lys | Ser | Asp | Ser | His | Leu | Gly | Asn | Glu | Leu | Lys | Leu | 110 |
| | | | | 110 | | | | | 115 | | | | | | 120 |
| Asn | Ile | Ser | Gln | Asp | Gln | Pro | Pro | Gly | Asp | Asn | Ala | Arg | Ser | Gln | 125 |
| | | | | 125 | | | | | 130 | | | | | | 135 |
| Val | Leu | Arg | Glu | Lys | Ile | Ile | Glu | Leu | Glu | Thr | Glu | Ile | Glu | Lys | 140 |
| | | | | 140 | | | | | 145 | | | | | | 150 |
| Phe | Lys | Ala | Glu | Asn | Ala | Ser | Leu | Ala | Lys | Leu | Arg | Ile | Glu | Arg | 155 |
| | | | | 155 | | | | | 160 | | | | | | 165 |
| Glu | Ser | Ala | Leu | Glu | Lys | Leu | Arg | Lys | Glu | Ile | Ala | Asp | Phe | Glu | 170 |
| | | | | 170 | | | | | 175 | | | | | | 180 |
| Gln | Gln | Lys | Ala | Lys | Glu | Leu | Ala | Arg | Ile | Glu | Glu | Phe | Lys | Lys | 185 |
| | | | | 185 | | | | | 190 | | | | | | 195 |
| Glu | Glu | Met | Arg | Lys | Leu | Gln | Lys | Glu | Arg | Lys | Val | Phe | Glu | Lys | 200 |
| | | | | 200 | | | | | 205 | | | | | | 210 |
| Tyr | Thr | Thr | Ala | Ala | Arg | Thr | Phe | Pro | Asp | Lys | Lys | Glu | Arg | Glu | 215 |
| | | | | 215 | | | | | 220 | | | | | | 225 |
| Glu | Ile | Gln | Thr | Leu | Lys | Gln | Gln | Ile | Ala | Asp | Leu | Arg | Glu | Asp | 230 |
| | | | | 230 | | | | | 235 | | | | | | 240 |
| Leu | Lys | Arg | Lys | Glu | Thr | Lys | Trp | Ser | Ser | Thr | His | Ser | Arg | Leu | 245 |
| | | | | 245 | | | | | 250 | | | | | | 255 |
| Arg | Ser | Gln | Ile | Gln | Met | Leu | Val | Arg | Glu | Asn | Thr | Asp | Leu | Arg | 260 |
| | | | | 260 | | | | | 265 | | | | | | 270 |
| Glu | Glu | Ile | Lys | Val | Met | Glu | Arg | Phe | Arg | Leu | Asp | Ala | Trp | Lys | 275 |
| | | | | 275 | | | | | 280 | | | | | | 285 |
| Arg | Ala | Glu | Ala | Ile | Glu | Ser | Ser | Leu | Glu | Val | Glu | Lys | Lys | Asp | 290 |
| | | | | 290 | | | | | 295 | | | | | | 300 |
| Lys | Leu | Ala | Asn | Thr | Ser | Val | Arg | Phe | Gln | Asn | Ser | Gln | Ile | Ser | 305 |
| | | | | 305 | | | | | 310 | | | | | | 315 |
| Ser | Gly | Thr | Gln | Val | Glu | Lys | Tyr | Lys | Lys | Asn | Tyr | Leu | Pro | Met | 320 |
| | | | | 320 | | | | | 325 | | | | | | 330 |
| Gln | Gly | Asn | Pro | Pro | Arg | Arg | Ser | Lys | Ser | Ala | Pro | Pro | Arg | Asp | 335 |
| | | | | 335 | | | | | 340 | | | | | | 345 |
| Leu | Gly | Asn | Leu | Asp | Lys | Gly | Gln | Ala | Ala | Ser | Pro | Arg | Glu | Pro | 350 |
| | | | | 350 | | | | | 355 | | | | | | 360 |
| Leu | Glu | Pro | Leu | Asn | Phe | Pro | Asp | Pro | Glu | Tyr | Lys | Glu | Glu | Glu | 365 |
| | | | | 365 | | | | | 370 | | | | | | 375 |
| Glu | Asp | Gln | Asp | Ile | Gln | Gly | Glu | Ile | Ser | His | Pro | Asp | Gly | Lys | 380 |
| | | | | 380 | | | | | 385 | | | | | | 390 |
| Val | Glu | Lys | Val | Tyr | Lys | Asn | Gly | Cys | Arg | Val | Ile | Leu | Phe | Pro | 395 |
| | | | | 395 | | | | | 400 | | | | | | 405 |
| Asn | Gly | Thr | Arg | Lys | Glu | Val | Ser | Ala | Asp | Gly | Lys | Thr | Ile | Thr | 410 |
| | | | | 410 | | | | | 415 | | | | | | 420 |
| Val | Thr | Phe | Phe | Asn | Gly | Asp | Val | Lys | Gln | Val | Met | Pro | Asp | Gln | 425 |
| | | | | 425 | | | | | 430 | | | | | | 435 |
| Arg | Val | Ile | Tyr | Tyr | Tyr | Ala | Ala | Ala | Gln | Thr | Thr | His | Thr | Thr | 440 |
| | | | | 440 | | | | | 445 | | | | | | 450 |
| Tyr | Pro | Glu | Gly | Leu | Glu | Val | Leu | His | Phe | Ser | Ser | Gly | Gln | Ile | 455 |
| | | | | 455 | | | | | 460 | | | | | | 465 |
| Glu | Lys | His | Tyr | Pro | Asp | Gly | Arg | Lys | Glu | Ile | Thr | Phe | Pro | Asp | 470 |
| | | | | 470 | | | | | 475 | | | | | | 480 |
| Gln | Thr | Val | Lys | Asn | Leu | Phe | Pro | Asp | Gly | Gln | Glu | Glu | Ser | Ile | 485 |
| | | | | 485 | | | | | 490 | | | | | | 495 |
| Phe | Pro | Asp | Gly | Thr | Ile | Val | Arg | Val | Gln | Arg | Asp | Gly | Asn | Lys | |

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Ile | Glu | Phe | 500 | Asn | Gly | Gln | Arg | 505 | Glu | Leu | His | Thr | Ala | Gln | 510 |
| | | | | 515 | | | | | 520 | | | | | | | 525 |
| Phe | Lys | Arg | Arg | 530 | Glu | Tyr | Pro | Asp | 535 | Thr | Val | Lys | Thr | Val | Tyr | 540 |
| | | | | 545 | | | | | 550 | | | | | | | 555 |
| Ala | Asn | Gly | His | 560 | Gln | Thr | Lys | Tyr | 565 | Arg | Ser | Gly | Arg | Ile | Arg | 565 |
| | | | | | | | | | | | | | | | | |
| Val | Lys | Asp | Lys | | Glu | Gly | Asn | Val | | Met | Asp | Thr | Glu | Leu | | |
| | | | | | | | | | | | | | | | | |

<210> 16

<211> 436

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1403289CD1

<400> 16

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|
| Met | Leu | Leu | Ser | Pro | Lys | Phe | Ser | Leu | Ser | Thr | Ile | His | Ile | Arg | | |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | | |
| Leu | Thr | Ala | Lys | Gly | Leu | Leu | Arg | Asn | Leu | Arg | Leu | Pro | Ser | Gly | | |
| | | | | 20 | | | | | 25 | | | | | 30 | | |
| Phe | Arg | Arg | Ser | Thr | Val | Val | Phe | His | Thr | Val | Glu | Lys | Ser | Arg | | |
| | | | | 35 | | | | | 40 | | | | | 45 | | |
| Gln | Lys | Asn | Pro | Arg | Ser | Leu | Cys | Ile | Gln | Pro | Gln | Thr | Ala | Pro | | |
| | | | | 50 | | | | | 55 | | | | | 60 | | |
| Asp | Ala | Leu | Pro | Pro | Glu | Lys | Thr | Leu | Glu | Leu | Thr | Gln | Tyr | Lys | | |
| | | | | 65 | | | | | 70 | | | | | 75 | | |
| Thr | Lys | Cys | Glu | Asn | Gln | Ser | Gly | Phe | Ile | Leu | Gln | Leu | Lys | Gln | | |
| | | | | 80 | | | | | 85 | | | | | 90 | | |
| Leu | Leu | Ala | Cys | Gly | Asn | Thr | Lys | Phe | Glu | Ala | Leu | Thr | Val | Val | | |
| | | | | 95 | | | | | 100 | | | | | 105 | | |
| Ile | Gln | His | Leu | Leu | Ser | Glu | Arg | Glu | Glu | Ala | Leu | Lys | Gln | His | | |
| | | | | 110 | | | | | 115 | | | | | 120 | | |
| Lys | Thr | Leu | Ser | Gln | Glu | Leu | Val | Asn | Leu | Arg | Gly | Glu | Leu | Val | | |
| | | | | 125 | | | | | 130 | | | | | 135 | | |
| Thr | Ala | Ser | Thr | Thr | Cys | Glu | Lys | Leu | Glu | Lys | Ala | Arg | Asn | Glu | | |
| | | | | 140 | | | | | 145 | | | | | 150 | | |
| Leu | Gln | Thr | Val | Tyr | Glu | Ala | Phe | Val | Gln | Gln | His | Gln | Ala | Glu | | |
| | | | | 155 | | | | | 160 | | | | | 165 | | |
| Lys | Thr | Glu | Arg | Glu | Asn | Arg | Leu | Lys | Glu | Phe | Tyr | Thr | Arg | Glu | | |
| | | | | 170 | | | | | 175 | | | | | 180 | | |
| Tyr | Glu | Lys | Leu | Arg | Asp | Thr | Tyr | Ile | Glu | Glu | Ala | Glu | Lys | Tyr | | |
| | | | | 185 | | | | | 190 | | | | | 195 | | |
| Lys | Met | Gln | Leu | Gln | Glu | Gln | Phe | Asp | Asn | Leu | Asn | Ala | Ala | His | | |
| | | | | 200 | | | | | 205 | | | | | 210 | | |
| Glu | Thr | Ser | Lys | Leu | Glu | Ile | Glu | Ala | Ser | His | Ser | Glu | Lys | Leu | | |
| | | | | 215 | | | | | 220 | | | | | 225 | | |
| Glu | Leu | Leu | Lys | Lys | Ala | Tyr | Glu | Ala | Ser | Leu | Ser | Glu | Ile | Lys | | |
| | | | | 230 | | | | | 235 | | | | | 240 | | |
| Lys | Gly | His | Glu | Ile | Glu | Lys | Lys | Ser | Leu | Glu | Asp | Leu | Leu | Ser | | |
| | | | | 245 | | | | | 250 | | | | | 255 | | |
| Glu | Lys | Gln | Glu | Ser | Leu | Glu | Lys | Gln | Ile | Asn | Asp | Leu | Lys | Ser | | |
| | | | | 260 | | | | | 265 | | | | | 270 | | |
| Glu | Asn | Asp | Ala | Leu | Asn | Glu | Lys | Leu | Lys | Ser | Glu | Glu | Gln | Lys | | |
| | | | | 275 | | | | | 280 | | | | | 285 | | |
| Arg | Arg | Ala | Arg | Glu | Lys | Ala | Asn | Leu | Lys | Asn | Pro | Gln | Ile | Met | | |
| | | | | 290 | | | | | 295 | | | | | 300 | | |
| Tyr | Leu | Glu | Gln | Glu | Leu | Glu | Ser | Leu | Lys | Ala | Val | Leu | Glu | Ile | | |
| | | | | 305 | | | | | 310 | | | | | 315 | | |
| Lys | Asn | Glu | Lys | Leu | His | Gln | Gln | Asp | Ile | Lys | Leu | Met | Lys | Met | | |
| | | | | 320 | | | | | 325 | | | | | 330 | | |
| Glu | Lys | Leu | Val | Asp | Asn | Asn | Thr | Ala | Leu | Val | Asp | Lys | Leu | Lys | | |
| | | | | 335 | | | | | 340 | | | | | 345 | | |
| Arg | Phe | Gln | Gln | Glu | Asn | Glu | Glu | Leu | Lys | Ala | Arg | Met | Asp | Lys | | |

| | | | | | |
|-----------------|-----|---------------------|-----|---------------------|-----|
| His Met Ala Ile | 350 | Ser Arg Gln Leu Ser | 355 | Thr Glu Gln Ala Val | 360 |
| | 365 | | 370 | | 375 |
| Gln Glu Ser Leu | 380 | Glu Lys Glu Ser Lys | 385 | Val Asn Lys Arg Leu | 390 |
| Met Glu Asn Glu | 395 | Leu Leu Trp Lys | 400 | Leu His Asn Gly Asp | 405 |
| Cys Ser Pro Lys | 410 | Arg Ser Pro Thr Ser | 415 | Ser Ala Ile Pro Leu | 420 |
| Ser Pro Arg Asn | 425 | Ser Gly Ser Leu His | 430 | Ser Pro Ser Ile Ser | 435 |

Arg

<210> 17
 <211> 363
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1607607CD1

<400> 17

| | | | |
|-----------------|-------------------------|-----------------|-----|
| Met Ala Ser Ala | Glu Leu Gln Gly Lys Tyr | Gln Lys Leu Ala | Gln |
| 1 | 5 | 10 | 15 |
| Glu Tyr Ser Lys | Leu Arg Ala Gln Asn Gln | Val Leu Lys Lys | Gly |
| | 20 | 25 | 30 |
| Val Val Asp Glu | Gln Ala Asn Ser Ala Ala | Leu Lys Glu Gln | Leu |
| | 35 | 40 | 45 |
| Lys Met Lys Asp | Gln Ser Leu Arg Lys Leu | Gln Gln Glu Met | Asp |
| | 50 | 55 | 60 |
| Ser Leu Thr Phe | Arg Asn Leu Gln Leu Ala | Lys Arg Val Glu | Leu |
| | 65 | 70 | 75 |
| Leu Gln Asp Glu | Leu Ala Leu Ser Glu Pro | Arg Gly Lys Lys | Asn |
| | 80 | 85 | 90 |
| Lys Lys Ser Gly | Glu Ser Ser Ser Gln Leu | Ser Gln Glu Gln | Lys |
| | 95 | 100 | 105 |
| Ser Val Phe Asp | Glu Asp Leu Gln Lys Lys | Ile Glu Glu Asn | Glu |
| | 110 | 115 | 120 |
| Arg Leu His Ile | Gln Phe Phe Glu Ala Asp | Glu Gln His Lys | His |
| | 125 | 130 | 135 |
| Val Glu Ala Glu | Leu Arg Ser Arg Leu Ala | Thr Leu Glu Thr | Glu |
| | 140 | 145 | 150 |
| Ala Ala Gln His | Gln Ala Val Val Asp Gly | Leu Thr Arg Lys | Tyr |
| | 155 | 160 | 165 |
| Met Glu Thr Ile | Glu Lys Leu Gln Asn Asp | Lys Ala Lys Leu | Glu |
| | 170 | 175 | 180 |
| Val Lys Ser Gln | Thr Leu Glu Lys Glu Ala | Lys Glu Cys Arg | Leu |
| | 185 | 190 | 195 |
| Arg Thr Glu Glu | Cys Gln Leu Gln Leu Lys | Thr Leu His Glu | Asp |
| | 200 | 205 | 210 |
| Leu Ser Gly Arg | Leu Glu Glu Ser Leu Ser | Ile Ile Asn Glu | Lys |
| | 215 | 220 | 225 |
| Val Pro Phe Asn | Asp Thr Lys Tyr Ser Gln | Tyr Asn Ala Leu | Asn |
| | 230 | 235 | 240 |
| Val Pro Leu His | Asn Arg Arg His Gln Leu | Lys Met Arg Asp | Ile |
| | 245 | 250 | 255 |
| Ala Gly Gln Ala | Leu Ala Phe Val Gln Asp | Leu Val Thr Ala | Leu |
| | 260 | 265 | 270 |
| Leu Asn Phe His | Thr Tyr Thr Glu Gln Arg | Ile Gln Ile Phe | Pro |
| | 275 | 280 | 285 |
| Val Asp Ser Ala | Ile Asp Thr Ile Ser Pro | Leu Asn Gln Lys | Phe |
| | 290 | 295 | 300 |
| Ser Gln Tyr Leu | His Glu Asn Ala Ser Tyr | Val Arg Pro Leu | Glu |
| | 305 | 310 | 315 |
| Glu Gly Met Leu | His Leu Phe Glu Ser Ile | Thr Glu Asp Thr | Val |

| | | | | | |
|-----------------|-----|---------------------|-----|-----------------|-----|
| Thr Val Leu Glu | 320 | Thr Thr Val Lys Leu | 325 | Thr Phe Ser Glu | 330 |
| | 335 | | 340 | | 345 |
| Leu Thr Ser Tyr | 350 | Ile Cys Phe Leu Arg | 355 | Ile Leu Pro Tyr | 360 |
| Leu Lys Arg | | | | | |

<210> 18
 <211> 247
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1660025CD1

<400> 18

| | | |
|---------------------|---------------------|---------------------|
| Met Gly Lys Arg Asp | Asn Arg Val Ala Tyr | Met Asn Pro Ile Ala |
| 1 | 5 | 10 |
| Met Ala Arg Ser Arg | Gly Pro Ile Gln Ser | Ser Gly Pro Thr Ile |
| | 20 | 25 |
| Gln Asp Tyr Leu Asn | Arg Pro Arg Pro Thr | Trp Glu Glu Val Lys |
| | 35 | 40 |
| Glu Gln Leu Glu Lys | Lys Lys Lys Gly Ser | Lys Ala Leu Ala Glu |
| | 50 | 55 |
| Phe Glu Glu Lys Met | Asn Glu Asn Trp Lys | Lys Glu Leu Glu Lys |
| | 65 | 70 |
| His Arg Glu Lys Leu | Leu Ser Gly Ser Glu | Ser Ser Ser Lys Lys |
| | 80 | 85 |
| Arg Gln Arg Lys Lys | Lys Glu Lys Lys Lys | Ser Gly Arg Tyr Ser |
| | 95 | 100 |
| Ser Ser Ser Ser Ser | Ser Ser Asp Ser Ser | Ser Ser Ser Ser Asp |
| | 110 | 115 |
| Ser Glu Asp Glu Asp | Lys Lys Gln Gly Lys | Arg Arg Lys Lys Lys |
| | 125 | 130 |
| Lys Asn Arg Ser His | Lys Ser Ser Glu Ser | Ser Met Ser Glu Thr |
| | 140 | 145 |
| Glu Ser Asp Ser Lys | Asp Ser Leu Lys Lys | Lys Lys Lys Ser Lys |
| | 155 | 160 |
| Asp Gly Thr Glu Lys | Glu Lys Asp Ile Lys | Gly Leu Ser Lys Lys |
| | 170 | 175 |
| Arg Lys Met Tyr Ser | Glu Asp Lys Pro Leu | Ser Ser Glu Ser Leu |
| | 185 | 190 |
| Ser Glu Ser Glu Tyr | Ile Glu Glu Val Arg | Ala Lys Lys Lys Lys |
| | 200 | 205 |
| Ser Ser Glu Glu Arg | Glu Lys Ala Thr Glu | Lys Thr Lys Lys Lys |
| | 215 | 220 |
| Lys Lys His Lys Lys | His Ser Lys Lys Lys | Lys Lys Lys Ala Ala |
| | 230 | 235 |
| Ser Ser Ser Pro Asp | Ser Pro | |
| | 245 | |

<210> 19
 <211> 441
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1796836CD1

<400> 19

| | | |
|---------------------|---------------------|---------------------|
| Met Asp Asp Asp Asp | Phe Gly Gly Phe Glu | Ala Ala Glu Thr Phe |
| 1 | 5 | 10 |
| Asp Gly Gly Ser Gly | Glu Thr Gln Thr Thr | Ser Pro Ala Ile Pro |
| | 20 | 25 |
| | | 30 |

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Trp | Ala | Ala | Phe | Pro | Ala | Val | Ser | Gly | Val | His | Leu | Ser | Pro | Ser |
| | | | | 35 | | | | | 40 | | | | | 45 |
| Ser | Pro | Glu | Ile | Val | Leu | Asp | Arg | Asp | His | Ser | Ser | Ser | Ile | Gly |
| | | | | 50 | | | | | 55 | | | | | 60 |
| Cys | Leu | Ser | Ser | Asp | Ala | Ile | Ile | Ser | Ser | Pro | Glu | Asn | Thr | His |
| | | | | 65 | | | | | 70 | | | | | 75 |
| Ala | Ala | Asn | Ser | Ile | Val | Ser | Gln | Thr | Ile | Pro | Lys | Ala | Gln | Ile |
| | | | | 80 | | | | | 85 | | | | | 90 |
| Gln | Gln | Ser | Thr | His | Thr | His | Leu | Asp | Ile | Ser | Leu | Phe | Pro | Leu |
| | | | | 95 | | | | | 100 | | | | | 105 |
| Gly | Leu | Thr | Asp | Glu | Lys | Ser | Asn | Gly | Thr | Ile | Ala | Leu | Val | Asp |
| | | | | 110 | | | | | 115 | | | | | 120 |
| Asp | Ser | Glu | Asp | Pro | Gly | Ala | Asn | Val | Ser | Asn | Ile | Gln | Leu | Gln |
| | | | | 125 | | | | | 130 | | | | | 135 |
| Gln | Lys | Ile | Ser | Ser | Leu | Glu | Ile | Lys | Leu | Lys | Val | Ser | Glu | Glu |
| | | | | 140 | | | | | 145 | | | | | 150 |
| Glu | Lys | Gln | Arg | Ile | Lys | Gln | Asp | Val | Glu | Ser | Leu | Met | Glu | Lys |
| | | | | 155 | | | | | 160 | | | | | 165 |
| His | Asn | Val | Leu | Glu | Lys | Gly | Phe | Leu | Lys | Glu | Lys | Glu | Gln | Glu |
| | | | | 170 | | | | | 175 | | | | | 180 |
| Ala | Ile | Ser | Phe | Gln | Asp | Arg | Tyr | Lys | Glu | Leu | Gln | Glu | Lys | His |
| | | | | 185 | | | | | 190 | | | | | 195 |
| Lys | Gln | Glu | Leu | Glu | Asp | Met | Arg | Lys | Ala | Gly | His | Glu | Ala | Leu |
| | | | | 200 | | | | | 205 | | | | | 210 |
| Ser | Ile | Ile | Val | Asp | Glu | Tyr | Lys | Ala | Leu | Leu | Gln | Ser | Ser | Val |
| | | | | 215 | | | | | 220 | | | | | 225 |
| Lys | Gln | Gln | Val | Glu | Ala | Ile | Glu | Lys | Gln | Tyr | Ile | Ser | Ala | Ile |
| | | | | 230 | | | | | 235 | | | | | 240 |
| Glu | Lys | Gln | Ala | His | Lys | Cys | Glu | Glu | Leu | Leu | Asn | Ala | Gln | His |
| | | | | 245 | | | | | 250 | | | | | 255 |
| Gln | Arg | Leu | Leu | Glu | Met | Leu | Asp | Thr | Glu | Lys | Glu | Leu | Leu | Lys |
| | | | | 260 | | | | | 265 | | | | | 270 |
| Glu | Lys | Ile | Lys | Glu | Ala | Leu | Ile | Gln | Gln | Ser | Gln | Glu | Gln | Lys |
| | | | | 275 | | | | | 280 | | | | | 285 |
| Glu | Ile | Leu | Glu | Lys | Cys | Leu | Glu | Glu | Glu | Arg | Gln | Arg | Asn | Lys |
| | | | | 290 | | | | | 295 | | | | | 300 |
| Glu | Ala | Leu | Val | Ser | Ala | Ala | Lys | Leu | Glu | Lys | Glu | Ala | Met | Lys |
| | | | | 305 | | | | | 310 | | | | | 315 |
| Asp | Ala | Val | Leu | Lys | Val | Val | Glu | Glu | Glu | Arg | Lys | Asn | Leu | Glu |
| | | | | 320 | | | | | 325 | | | | | 330 |
| Lys | Ala | His | Ala | Glu | Glu | Arg | Glu | Leu | Trp | Lys | Thr | Glu | His | Ala |
| | | | | 335 | | | | | 340 | | | | | 345 |
| Lys | Asp | Gln | Glu | Lys | Val | Ser | Gln | Glu | Ile | Gln | Lys | Ala | Ile | Gln |
| | | | | 350 | | | | | 355 | | | | | 360 |
| Glu | Gln | Arg | Lys | Ile | Ser | Gln | Glu | Thr | Val | Lys | Ala | Ala | Ile | Ile |
| | | | | 365 | | | | | 370 | | | | | 375 |
| Glu | Glu | Gln | Lys | Arg | Ser | Glu | Lys | Ala | Val | Glu | Glu | Ala | Val | Lys |
| | | | | 380 | | | | | 385 | | | | | 390 |
| Arg | Thr | Arg | Asp | Glu | Leu | Ile | Glu | Tyr | Ile | Lys | Glu | Gln | Lys | Arg |
| | | | | 395 | | | | | 400 | | | | | 405 |
| Leu | Asp | Gln | Val | Ile | Arg | Gln | Arg | Ser | Leu | Ser | Ser | Leu | Glu | Leu |
| | | | | 410 | | | | | 415 | | | | | 420 |
| Phe | Leu | Ser | Cys | Ala | Gln | Lys | Gln | Leu | Ser | Ala | Leu | Ile | Ala | Thr |
| | | | | 425 | | | | | 430 | | | | | 435 |
| Glu | Pro | Val | Asp | Ile | Glu | | | | | | | | | |
| | | | | 440 | | | | | | | | | | |

<210> 20

<211> 183

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2880670CD1

<400> 20


```

Met Ala Ala Gln Arg Gly Met Pro Ser Ser Ala Val Arg Val Leu
 1      5      10
Glu Glu Ala Leu Gly Met Gly Leu Thr Ala Ala Gly Asp Ala Arg
 20      25      30
Asp Thr Ala Asp Ala Val Ala Ala Glu Gly Ala Tyr Tyr Leu Glu
 35      40      45
Gln Val Thr Ile Thr Glu Ala Ser Glu Asp Asp Tyr Glu Tyr Glu
 50      55      60
Glu Ile Pro Asp Asp Asn Phe Ser Ile Pro Glu Gly Glu Glu Asp
 65      70      75
Leu Ala Lys Ala Ile Gln Met Ala Gln Glu Gln Ala Thr Asp Thr
 80      85      90
Glu Ile Leu Glu Arg Lys Thr Val Leu Pro Ser Lys His Ala Val
 95      100     105
Pro Glu Val Ile Glu Asp Phe Leu Cys Asn Phe Leu Ile Lys Met
 110     115     120
Gly Met Thr Arg Thr Leu Asp Cys Phe Gln Ser Glu Trp Tyr Glu
 125     130     135
Leu Ile Gln Lys Gly Val Thr Glu Leu Arg Thr Val Gly Asn Val
 140     145     150
Pro Asp Val Tyr Thr Gln Ile Met Leu Leu Glu Asn Glu Asn Lys
 155     160     165
Asn Leu Lys Lys Asp Leu Lys His Tyr Lys Gln Ala Ala Glu Tyr
 170     175     180
Val Ile Phe

```

<210> 21
 <211> 212
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2913976CD1

```

<400> 21
Met Glu Glu Glu Leu Pro Pro Pro Pro Ala Glu Pro Val Glu Lys
 1      5      10      15
Gly Ala Ser Thr Asp Ile Cys Ala Phe Cys His Lys Thr Val Phe
 20      25      30
Pro Arg Glu Leu Ala Val Glu Ala Met Lys Arg Gln Tyr His Ala
 35      40      45
Gln Cys Phe Thr Cys Arg Thr Cys Arg Arg Gln Leu Ala Gly Gln
 50      55      60
Ser Phe Tyr Gln Lys Asp Gly Arg Pro Leu Cys Glu Pro Cys Tyr
 65      70      75
Gln Asp Thr Leu Glu Arg Cys Gly Lys Cys Gly Glu Val Val Arg
 80      85      90
Asp His Ile Ile Arg Ala Leu Gly Gln Ala Phe His Pro Ser Cys
 95      100     105
Phe Thr Cys Val Thr Cys Ala Arg Cys Ile Gly Asp Glu Ser Phe
 110     115     120
Ala Leu Gly Ser Gln Asn Glu Val Tyr Cys Leu Asp Asp Phe Tyr
 125     130     135
Arg Lys Phe Ala Pro Val Cys Ser Ile Cys Glu Asn Pro Ile Ile
 140     145     150
Pro Arg Asp Gly Lys Asp Ala Phe Lys Ile Glu Cys Met Gly Arg
 155     160     165
Asn Phe His Glu Asn Cys Tyr Arg Cys Glu Asp Cys Arg Ile Leu
 170     175     180
Leu Ser Val Glu Pro Thr Asp Gln Gly Cys Tyr Pro Leu Asn Asn
 185     190     195
His Leu Phe Cys Lys Pro Cys His Val Lys Arg Ser Ala Ala Gly
 200     205     210
Cys Cys

```

<210> 22
 <211> 227
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3092084CD1

<400> 22
 Met Gly Gly Thr Thr Ser Thr Arg Arg Val Thr Phe Glu Ala Asp
 1 5 10 15
 Glu Asn Glu Asn Ile Thr Val Val Lys Gly Ile Arg Leu Ser Glu
 20 25 30
 Asn Val Ile Asp Arg Met Lys Glu Ser Ser Pro Ser Gly Ser Lys
 35 40 45
 Ser Gln Arg Tyr Ser Gly Ala Tyr Gly Ala Ser Val Ser Asp Glu
 50 55 60
 Glu Leu Lys Arg Arg Val Ala Glu Glu Leu Ala Leu Glu Gln Ala
 65 70 75
 Lys Lys Glu Ser Glu Asp Gln Lys Arg Leu Lys Gln Ala Lys Glu
 80 85 90
 Leu Asp Arg Glu Arg Ala Ala Ala Asn Glu Gln Leu Thr Arg Ala
 95 100 105
 Ile Leu Arg Glu Arg Ile Cys Ser Glu Glu Glu Arg Ala Lys Ala
 110 115 120
 Lys His Leu Ala Arg Gln Leu Glu Glu Lys Asp Arg Val Leu Lys
 125 130 135
 Lys Gln Asp Ala Phe Tyr Lys Glu Gln Leu Ala Arg Leu Glu Glu
 140 145 150
 Arg Ser Ser Glu Phe Tyr Arg Val Thr Thr Glu Gln Tyr Gln Lys
 155 160 165
 Ala Ala Glu Glu Val Glu Ala Lys Phe Lys Arg Tyr Glu Ser His
 170 175 180
 Pro Val Cys Ala Asp Leu Gln Ala Lys Ile Leu Gln Cys Tyr Arg
 185 190 195
 Glu Asn Thr His Gln Thr Leu Lys Cys Ser Ala Leu Ala Thr Gln
 200 205 210
 Tyr Met His Cys Val Asn His Ala Lys Gln Ser Met Leu Glu Lys
 215 220 225
 Gly Gly

<210> 23
 <211> 490
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3882482CD1

<400> 23
 Met Asn Leu Ala Glu Ile Cys Asp Asn Ala Lys Lys Gly Arg Glu
 1 5 10 15
 Tyr Ala Leu Leu Gly Asn Tyr Asp Ser Ser Met Val Tyr Tyr Gln
 20 25 30
 Gly Val Met Gln Gln Ile Gln Arg His Cys Gln Ser Val Arg Asp
 35 40 45
 Pro Ala Ile Lys Gly Lys Trp Gln Gln Val Arg Gln Glu Leu Leu
 50 55 60
 Glu Glu Tyr Glu Gln Val Lys Ser Ile Val Ser Thr Leu Glu Ser
 65 70 75
 Phe Lys Ile Asp Lys Pro Pro Asp Phe Pro Val Ser Cys Gln Asp
 80 85 90
 Glu Pro Phe Arg Asp Pro Ala Val Trp Pro Pro Pro Val Pro Ala
 95 100 105

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Glu | His | Arg | Ala | Pro | Pro | Gln | Ile | Arg | Arg | Pro | Asn | Arg | Glu | Val |
| | | | | 110 | | | | | 115 | | | | | 120 |
| Arg | Pro | Leu | Arg | Lys | Glu | Met | Ala | Gly | Val | Gly | Ala | Arg | Gly | Pro |
| | | | | 125 | | | | | 130 | | | | | 135 |
| Val | Gly | Arg | Ala | His | Pro | Ile | Ser | Lys | Ser | Glu | Lys | Pro | Ser | Thr |
| | | | | 140 | | | | | 145 | | | | | 150 |
| Ser | Arg | Asp | Lys | Asp | Tyr | Arg | Ala | Arg | Gly | Arg | Asp | Asp | Lys | Gly |
| | | | | 155 | | | | | 160 | | | | | 165 |
| Arg | Lys | Asn | Met | Gln | Asp | Gly | Ala | Ser | Asn | Gly | Glu | Met | Pro | Lys |
| | | | | 170 | | | | | 175 | | | | | 180 |
| Phe | Asp | Gly | Ala | Gly | Tyr | Asp | Lys | Asp | Leu | Val | Glu | Ala | Leu | Glu |
| | | | | 185 | | | | | 190 | | | | | 195 |
| Arg | Asp | Ile | Val | Ser | Arg | Asn | Pro | Ser | Ile | His | Trp | Asp | Asp | Ile |
| | | | | 200 | | | | | 205 | | | | | 210 |
| Ala | Asp | Leu | Glu | Glu | Ala | Lys | Lys | Leu | Leu | Arg | Glu | Ala | Val | Val |
| | | | | 215 | | | | | 220 | | | | | 225 |
| Leu | Pro | Met | Trp | Met | Pro | Asp | Phe | Phe | Lys | Gly | Ile | Arg | Arg | Pro |
| | | | | 230 | | | | | 235 | | | | | 240 |
| Trp | Lys | Gly | Val | Leu | Met | Val | Gly | Pro | Pro | Gly | Thr | Gly | Lys | Thr |
| | | | | 245 | | | | | 250 | | | | | 255 |
| Met | Leu | Ala | Lys | Ala | Val | Ala | Thr | Glu | Cys | Gly | Thr | Thr | Phe | Phe |
| | | | | 260 | | | | | 265 | | | | | 270 |
| Asn | Val | Ser | Ser | Ser | Thr | Leu | Thr | Ser | Lys | Tyr | Arg | Gly | Glu | Ser |
| | | | | 275 | | | | | 280 | | | | | 285 |
| Glu | Lys | Leu | Val | Arg | Leu | Leu | Phe | Glu | Met | Ala | Arg | Phe | Tyr | Ala |
| | | | | 290 | | | | | 295 | | | | | 300 |
| Pro | Thr | Thr | Ile | Phe | Ile | Asp | Glu | Ile | Asp | Ser | Ile | Cys | Ser | Arg |
| | | | | 305 | | | | | 310 | | | | | 315 |
| Arg | Gly | Thr | Ser | Asp | Glu | His | Glu | Ala | Ser | Arg | Arg | Val | Lys | Ser |
| | | | | 320 | | | | | 325 | | | | | 330 |
| Glu | Leu | Leu | Ile | Gln | Met | Asp | Gly | Val | Gly | Gly | Ala | Leu | Glu | Asn |
| | | | | 335 | | | | | 340 | | | | | 345 |
| Asp | Asp | Pro | Ser | Lys | Met | Val | Met | Val | Leu | Ala | Ala | Thr | Asn | Phe |
| | | | | 350 | | | | | 355 | | | | | 360 |
| Pro | Trp | Asp | Ile | Asp | Glu | Ala | Leu | Arg | Arg | Arg | Leu | Glu | Lys | Arg |
| | | | | 365 | | | | | 370 | | | | | 375 |
| Ile | Tyr | Ile | Pro | Leu | Pro | Thr | Ala | Lys | Gly | Arg | Ala | Glu | Leu | Leu |
| | | | | 380 | | | | | 385 | | | | | 390 |
| Lys | Ile | Asn | Leu | Arg | Glu | Val | Glu | Leu | Asp | Pro | Asp | Ile | Gln | Leu |
| | | | | 395 | | | | | 400 | | | | | 405 |
| Glu | Asp | Ile | Ala | Glu | Lys | Ile | Glu | Gly | Tyr | Ser | Gly | Ala | Asp | Ile |
| | | | | 410 | | | | | 415 | | | | | 420 |
| Thr | Asn | Val | Cys | Arg | Asp | Ala | Ser | Leu | Met | Ala | Met | Arg | Arg | Arg |
| | | | | 425 | | | | | 430 | | | | | 435 |
| Ile | Asn | Gly | Leu | Ser | Pro | Glu | Glu | Ile | Arg | Ala | Leu | Ser | Lys | Glu |
| | | | | 440 | | | | | 445 | | | | | 450 |
| Glu | Leu | Gln | Met | Pro | Val | Thr | Lys | Gly | Asp | Phe | Glu | Leu | Ala | Leu |
| | | | | 455 | | | | | 460 | | | | | 465 |
| Lys | Lys | Ile | Ala | Lys | Ser | Val | Ser | Ala | Ala | Asp | Leu | Glu | Lys | Tyr |
| | | | | 470 | | | | | 475 | | | | | 480 |
| Glu | Lys | Trp | Met | Val | Glu | Phe | Gly | Ser | Ala | | | | | |
| | | | | 485 | | | | | 490 | | | | | |

<210> 24

<211> 133

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4933451CD1

<400> 24

Met Ala Ala Arg Thr Val Ile Ile Asp His Gly Ser Gly Phe Leu

1 5 10 15

Lys Ala Gly Thr Ala Gly Trp Asn Glu Pro Gln Met Val Phe Pro

20 25 30

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | Ile | Val | Asn | Tyr | Leu | Pro | Cys | Lys | Glu | Asn | Pro | Gly | Pro | Ser |
| | | | | 35 | | | | | 40 | | | | | 45 |
| Tyr | Ala | Arg | Lys | Arg | Val | Ser | Leu | Gly | Ile | Asp | Ile | Cys | His | Pro |
| | | | | 50 | | | | | 55 | | | | | 60 |
| Asp | Thr | Phe | Ser | Tyr | Pro | Ile | Glu | Arg | Gly | Arg | Ile | Leu | Asn | Trp |
| | | | | 65 | | | | | 70 | | | | | 75 |
| Glu | Gly | Val | Gln | Tyr | Leu | Trp | Ser | Phe | Val | Leu | Glu | Asn | His | Arg |
| | | | | 80 | | | | | 85 | | | | | 90 |
| Arg | Glu | Gln | Glu | Val | Pro | Pro | Val | Ile | Ile | Thr | Glu | Thr | Pro | Leu |
| | | | | 95 | | | | | 100 | | | | | 105 |
| Arg | Glu | Pro | Ala | Asp | Arg | Lys | Lys | Met | Ser | Ser | Leu | Glu | Thr | Leu |
| | | | | 110 | | | | | 115 | | | | | 120 |
| Gln | Gly | Thr | Val | Phe | Pro | Gly | Trp | Pro | Ile | Ile | Gly | Val | | |
| | | | | 125 | | | | | 130 | | | | | |

<210> 25

<211> 912

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5043904CD1

<400> 25

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ser | Asp | Ala | Gln | Gly | Ser | Tyr | Lys | Leu | Asp | Glu | Ala | Gln | Ala |
| 1 | | | | 5 | | | | | 10 | | | | | 15 |
| Val | Leu | Arg | Glu | Thr | Lys | Ala | Ile | Lys | Lys | Ala | Ile | Thr | Cys | Gly |
| | | | | 20 | | | | | 25 | | | | | 30 |
| Glu | Lys | Glu | Lys | Gln | Asp | Leu | Ile | Lys | Ser | Leu | Ala | Met | Leu | Lys |
| | | | | 35 | | | | | 40 | | | | | 45 |
| Asp | Gly | Phe | Cys | Thr | Asp | Arg | Gly | Ser | His | Ser | Asp | Leu | Trp | Ser |
| | | | | 50 | | | | | 55 | | | | | 60 |
| Ser | Ser | Ser | Ser | Leu | Glu | Ser | Ser | Ser | Phe | Pro | Leu | Pro | Lys | Gln |
| | | | | 65 | | | | | 70 | | | | | 75 |
| Tyr | Leu | Asp | Val | Ser | Ser | Gln | Thr | Asp | Ile | Ser | Gly | Ser | Phe | Gly |
| | | | | 80 | | | | | 85 | | | | | 90 |
| Ile | Asn | Ser | Asn | Asn | Gln | Leu | Ala | Glu | Lys | Val | Arg | Leu | Arg | Leu |
| | | | | 95 | | | | | 100 | | | | | 105 |
| Arg | Tyr | Glu | Glu | Ala | Lys | Arg | Arg | Ile | Ala | Asn | Leu | Lys | Ile | Gln |
| | | | | 110 | | | | | 115 | | | | | 120 |
| Leu | Ala | Lys | Leu | Asp | Ser | Glu | Ala | Trp | Pro | Gly | Val | Leu | Asp | Ser |
| | | | | 125 | | | | | 130 | | | | | 135 |
| Glu | Arg | Asp | Arg | Leu | Ile | Leu | Ile | Asn | Glu | Lys | Glu | Glu | Leu | Leu |
| | | | | 140 | | | | | 145 | | | | | 150 |
| Lys | Glu | Met | Arg | Phe | Ile | Ser | Pro | Arg | Lys | Trp | Thr | Gln | Gly | Glu |
| | | | | 155 | | | | | 160 | | | | | 165 |
| Val | Glu | Gln | Leu | Glu | Met | Ala | Arg | Lys | Arg | Leu | Glu | Lys | Asp | Leu |
| | | | | 170 | | | | | 175 | | | | | 180 |
| Gln | Ala | Ala | Arg | Asp | Thr | Gln | Ser | Lys | Ala | Leu | Thr | Glu | Arg | Leu |
| | | | | 185 | | | | | 190 | | | | | 195 |
| Lys | Leu | Asn | Ser | Lys | Arg | Asn | Gln | Leu | Val | Arg | Glu | Leu | Glu | Glu |
| | | | | 200 | | | | | 205 | | | | | 210 |
| Ala | Thr | Arg | Gln | Val | Ala | Thr | Leu | His | Ser | Gln | Leu | Lys | Ser | Leu |
| | | | | 215 | | | | | 220 | | | | | 225 |
| Ser | Ser | Ser | Met | Gln | Ser | Leu | Ser | Ser | Gly | Ser | Ser | Pro | Gly | Ser |
| | | | | 230 | | | | | 235 | | | | | 240 |
| Leu | Thr | Ser | Ser | Arg | Gly | Ser | Leu | Val | Ala | Ser | Ser | Leu | Asp | Ser |
| | | | | 245 | | | | | 250 | | | | | 255 |
| Ser | Thr | Ser | Ala | Ser | Phe | Thr | Asp | Leu | Tyr | Tyr | Asp | Pro | Phe | Glu |
| | | | | 260 | | | | | 265 | | | | | 270 |
| Gln | Leu | Asp | Ser | Glu | Leu | Gln | Ser | Lys | Val | Glu | Phe | Leu | Leu | Leu |
| | | | | 275 | | | | | 280 | | | | | 285 |
| Glu | Gly | Ala | Thr | Gly | Phe | Arg | Pro | Ser | Gly | Cys | Ile | Thr | Thr | Ile |
| | | | | 290 | | | | | 295 | | | | | 300 |
| His | Glu | Asp | Glu | Val | Ala | Lys | Thr | Gln | Lys | Ala | Glu | Gly | Gly | Gly |
| | | | | 305 | | | | | 310 | | | | | 315 |

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Arg | Leu | Gln | Ala | Leu | Arg | Ser | Leu | Ser | Gly | Thr | Pro | Lys | Ser | Met |
| | | | | 320 | | | | | 325 | | | | | 330 |
| Thr | Ser | Leu | Ser | Pro | Arg | Ser | Ser | Leu | Ser | Ser | Pro | Ser | Pro | Pro |
| | | | | 335 | | | | | 340 | | | | | 345 |
| Cys | Ser | Pro | Leu | Met | Ala | Asp | Pro | Leu | Leu | Ala | Gly | Asp | Ala | Phe |
| | | | | 350 | | | | | 355 | | | | | 360 |
| Leu | Asn | Ser | Leu | Glu | Phe | Glu | Asp | Pro | Glu | Leu | Ser | Ala | Thr | Leu |
| | | | | 365 | | | | | 370 | | | | | 375 |
| Cys | Glu | Leu | Ser | Leu | Gly | Asn | Ser | Ala | Gln | Glu | Arg | Tyr | Arg | Leu |
| | | | | 380 | | | | | 385 | | | | | 390 |
| Glu | Glu | Pro | Gly | Thr | Glu | Gly | Lys | Gln | Leu | Gly | Gln | Ala | Val | Asn |
| | | | | 395 | | | | | 400 | | | | | 405 |
| Thr | Ala | Gln | Gly | Cys | Gly | Leu | Lys | Val | Ala | Cys | Val | Ser | Ala | Ala |
| | | | | 410 | | | | | 415 | | | | | 420 |
| Val | Ser | Asp | Glu | Ser | Val | Ala | Gly | Asp | Ser | Gly | Val | Tyr | Glu | Ala |
| | | | | 425 | | | | | 430 | | | | | 435 |
| Ser | Val | Gln | Arg | Leu | Gly | Ala | Ser | Glu | Ala | Ala | Ala | Phe | Asp | Ser |
| | | | | 440 | | | | | 445 | | | | | 450 |
| Asp | Glu | Ser | Glu | Ala | Val | Gly | Ala | Thr | Arg | Ile | Gln | Ile | Ala | Leu |
| | | | | 455 | | | | | 460 | | | | | 465 |
| Lys | Tyr | Asp | Glu | Lys | Asn | Lys | Gln | Phe | Ala | Ile | Leu | Ile | Ile | Gln |
| | | | | 470 | | | | | 475 | | | | | 480 |
| Leu | Ser | Asn | Leu | Ser | Ala | Leu | Leu | Gln | Gln | Gln | Asp | Gln | Lys | Val |
| | | | | 485 | | | | | 490 | | | | | 495 |
| Asn | Ile | Arg | Val | Ala | Val | Leu | Pro | Cys | Ser | Glu | Ser | Thr | Thr | Cys |
| | | | | 500 | | | | | 505 | | | | | 510 |
| Leu | Phe | Arg | Thr | Arg | Pro | Leu | Asp | Ala | Ser | Asp | Thr | Leu | Val | Phe |
| | | | | 515 | | | | | 520 | | | | | 525 |
| Asn | Glu | Val | Phe | Trp | Val | Ser | Met | Ser | Tyr | Pro | Ala | Leu | His | Gln |
| | | | | 530 | | | | | 535 | | | | | 540 |
| Lys | Thr | Leu | Arg | Val | Asp | Val | Cys | Thr | Thr | Asp | Arg | Ser | His | Leu |
| | | | | 545 | | | | | 550 | | | | | 555 |
| Glu | Glu | Cys | Leu | Gly | Gly | Ala | Gln | Ile | Ser | Leu | Ala | Glu | Val | Cys |
| | | | | 560 | | | | | 565 | | | | | 570 |
| Arg | Ser | Gly | Glu | Arg | Ser | Thr | Arg | Trp | Tyr | Asn | Leu | Leu | Ser | Tyr |
| | | | | 575 | | | | | 580 | | | | | 585 |
| Lys | Tyr | Leu | Lys | Lys | Gln | Ser | Arg | Glu | Leu | Lys | Pro | Val | Gly | Val |
| | | | | 590 | | | | | 595 | | | | | 600 |
| Met | Ala | Pro | Ala | Ser | Gly | Pro | Ala | Ser | Thr | Asp | Ala | Val | Ser | Ala |
| | | | | 605 | | | | | 610 | | | | | 615 |
| Leu | Leu | Glu | Gln | Thr | Ala | Val | Glu | Leu | Glu | Lys | Arg | Gln | Glu | Gly |
| | | | | 620 | | | | | 625 | | | | | 630 |
| Arg | Ser | Ser | Thr | Gln | Thr | Leu | Glu | Asp | Ser | Trp | Arg | Tyr | Glu | Glu |
| | | | | 635 | | | | | 640 | | | | | 645 |
| Thr | Ser | Glu | Asn | Glu | Ala | Val | Ala | Glu | Glu | Glu | Glu | Glu | Glu | Val |
| | | | | 650 | | | | | 655 | | | | | 660 |
| Glu | Glu | Glu | Glu | Gly | Glu | Glu | Asp | Val | Phe | Thr | Glu | Lys | Ala | Ser |
| | | | | 665 | | | | | 670 | | | | | 675 |
| Pro | Asp | Met | Asp | Gly | Tyr | Pro | Ala | Leu | Lys | Val | Asp | Lys | Glu | Thr |
| | | | | 680 | | | | | 685 | | | | | 690 |
| Asn | Thr | Glu | Thr | Pro | Ala | Pro | Ser | Pro | Thr | Val | Val | Arg | Pro | Lys |
| | | | | 695 | | | | | 700 | | | | | 705 |
| Asp | Arg | Arg | Val | Gly | Thr | Pro | Ser | Gln | Gly | Pro | Phe | Leu | Arg | Gly |
| | | | | 710 | | | | | 715 | | | | | 720 |
| Ser | Thr | Ile | Ile | Arg | Ser | Lys | Thr | Phe | Ser | Pro | Gly | Pro | Gln | Ser |
| | | | | 725 | | | | | 730 | | | | | 735 |
| Gln | Tyr | Val | Cys | Arg | Leu | Asn | Arg | Ser | Asp | Ser | Asp | Ser | Ser | Thr |
| | | | | 740 | | | | | 745 | | | | | 750 |
| Leu | Ser | Lys | Lys | Pro | Pro | Phe | Val | Arg | Asn | Ser | Leu | Glu | Arg | Arg |
| | | | | 755 | | | | | 760 | | | | | 765 |
| Ser | Val | Arg | Met | Lys | Arg | Pro | Ser | Ser | Val | Lys | Ser | Leu | Arg | Ser |
| | | | | 770 | | | | | 775 | | | | | 780 |
| Glu | Arg | Leu | Ile | Arg | Thr | Ser | Leu | Asp | Leu | Glu | Leu | Asp | Leu | Gln |
| | | | | 785 | | | | | 790 | | | | | 795 |
| Ala | Thr | Arg | Thr | Trp | His | Ser | Gln | Leu | Thr | Gln | Glu | Ile | Ser | Val |
| | | | | 800 | | | | | 805 | | | | | 810 |
| Leu | Lys | Glu | Leu | Lys | Glu | Gln | Leu | Glu | Gln | Ala | Lys | Ser | His | Gly |

| | | | | | |
|-----------------|-----|---------------------|-----|---------------------|-----|
| Glu Lys Glu Leu | 815 | Gln Trp Leu Arg | 820 | Glu Asp Glu Arg Phe | 825 |
| Pro | 830 | Met | 835 | Ala Glu | 840 |
| Leu Leu Leu Arg | 845 | Gln Met Asp Arg | 850 | Ala | 855 |
| His Lys Gly Glu | 860 | Met Met Arg Ala Ala | 865 | Ala | 870 |
| Lys Asp Val His | 875 | Ser Cys Lys Glu Pro | 880 | Pro | 885 |
| Glu Val Gln Ser | 890 | Ala Phe Phe Thr Arg | 895 | Pro | 900 |
| Arg Met Asn Ile | 905 | Asp Asp Val | 910 | | |

<210> 26

<211> 1076

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5202390CD1

<400> 26

| | | | | | | | | | | | |
|-----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met Lys Gln Tyr | Ala | Ser | Pro | Met | Pro | Thr | Gln | Thr | Asp | Val | Lys |
| 1 | 5 | | | | | 10 | | | | | 15 |
| Leu Lys Phe Lys | Pro | Leu | Ser | Lys | Lys | Val | Val | Ser | Ala | Ala | Leu |
| | 20 | | | | | 25 | | | | | 30 |
| Gln Phe Ser Leu | Ser | Cys | Ile | Phe | Leu | Arg | Glu | Gly | Lys | Ala | Thr |
| | 35 | | | | | 40 | | | | | 45 |
| Asp Glu Asp Met | Gln | Ser | Leu | Ala | Ser | Leu | Met | Ser | Met | Lys | Gln |
| | 50 | | | | | 55 | | | | | 60 |
| Ala Asp Ile Gly | Asn | Leu | Asp | Asp | Phe | Glu | Glu | Asp | Asn | Glu | Asp |
| | 65 | | | | | 70 | | | | | 75 |
| Asp Asp Glu Asn | Arg | Val | Asn | Gln | Glu | Glu | Lys | Ala | Ala | Lys | Ile |
| | 80 | | | | | 85 | | | | | 90 |
| Thr Glu Leu Ile | Asn | Lys | Leu | Asn | Phe | Leu | Asp | Glu | Ala | Glu | Lys |
| | 95 | | | | | 100 | | | | | 105 |
| Asp Leu Ala Thr | Val | Asn | Ser | Asn | Pro | Phe | Asp | Asp | Pro | Asp | Ala |
| | 110 | | | | | 115 | | | | | 120 |
| Ala Glu Leu Asn | Pro | Phe | Gly | Asp | Pro | Asp | Ser | Glu | Glu | Pro | Ile |
| | 125 | | | | | 130 | | | | | 135 |
| Thr Glu Thr Ala | Ser | Pro | Arg | Lys | Thr | Glu | Asp | Ser | Phe | Tyr | Asn |
| | 140 | | | | | 145 | | | | | 150 |
| Asn Ser Tyr Asn | Pro | Phe | Lys | Glu | Val | Gln | Thr | Pro | Gln | Tyr | Leu |
| | 155 | | | | | 160 | | | | | 165 |
| Asn Pro Phe Asp | Glu | Pro | Glu | Ala | Phe | Val | Thr | Ile | Lys | Asp | Ser |
| | 170 | | | | | 175 | | | | | 180 |
| Pro Pro Gln Ser | Thr | Lys | Arg | Lys | Asn | Ile | Arg | Pro | Val | Asp | Met |
| | 185 | | | | | 190 | | | | | 195 |
| Ser Lys Tyr Leu | Tyr | Ala | Asp | Ser | Ser | Lys | Thr | Glu | Glu | Glu | Glu |
| | 200 | | | | | 205 | | | | | 210 |
| Leu Asp Glu Ser | Asn | Pro | Phe | Tyr | Glu | Pro | Lys | Ser | Thr | Pro | Pro |
| | 215 | | | | | 220 | | | | | 225 |
| Pro Asn Asn Leu | Val | Asn | Pro | Val | Gln | Glu | Leu | Glu | Thr | Glu | Arg |
| | 230 | | | | | 235 | | | | | 240 |
| Arg Val Lys Arg | Lys | Ala | Pro | Ala | Pro | Pro | Val | Leu | Ser | Pro | Lys |
| | 245 | | | | | 250 | | | | | 255 |
| Thr Gly Val Leu | Asn | Glu | Asn | Thr | Val | Ser | Ala | Gly | Lys | Asp | Leu |
| | 260 | | | | | 265 | | | | | 270 |
| Ser Thr Ser Pro | Lys | Pro | Ser | Pro | Ile | Pro | Ser | Pro | Val | Leu | Gly |
| | 275 | | | | | 280 | | | | | 285 |
| Arg Lys Pro Asn | Ala | Ser | Gln | Ser | Leu | Leu | Val | Trp | Cys | Lys | Glu |
| | 290 | | | | | 295 | | | | | 300 |
| Val Thr Lys Asn | Tyr | Arg | Gly | Val | Lys | Ile | Thr | Asn | Phe | Thr | Thr |
| | 305 | | | | | 310 | | | | | 315 |
| Ser Trp Arg Asn | Gly | Leu | Ser | Phe | Cys | Ala | Ile | Leu | His | His | Phe |

| | | | | | |
|-----------------|---------------------|---------------------|-----|--|-----|
| | 320 | | 325 | | 330 |
| Arg Pro Asp Leu | Ile Asp Tyr Lys Ser | Leu Asn Pro Gln Asp | Ile | | |
| | 335 | | 340 | | 345 |
| Lys Glu Asn Asn | Lys Lys Ala Tyr Asp | Gly Phe Ala Ser Ile | Gly | | |
| | 350 | | 355 | | 360 |
| Ile Ser Arg Leu | Leu Glu Pro Ser Asp | Met Val Leu Leu Ala | Ile | | |
| | 365 | | 370 | | 375 |
| Pro Asp Lys Leu | Thr Val Met Thr Tyr | Leu Tyr Gln Ile Arg | Ala | | |
| | 380 | | 385 | | 390 |
| His Phe Ser Gly | Gln Glu Leu Asn Val | Val Gln Ile Glu Glu | Asn | | |
| | 395 | | 400 | | 405 |
| Ser Ser Lys Ser | Thr Tyr Lys Val Gly | Asn Tyr Glu Thr Asp | Thr | | |
| | 410 | | 415 | | 420 |
| Asn Ser Ser Val | Asp Gln Glu Lys Phe | Tyr Ala Glu Leu Ser | Asp | | |
| | 425 | | 430 | | 435 |
| Leu Lys Arg Glu | Pro Glu Leu Gln Gln | Pro Ile Ser Gly Ala | Val | | |
| | 440 | | 445 | | 450 |
| Asp Phe Leu Ser | Gln Asp Asp Ser Val | Phe Val Asn Asp Ser | Gly | | |
| | 455 | | 460 | | 465 |
| Val Gly Glu Ser | Glu Ser Glu His Gln | Thr Pro Asp Asp His | Leu | | |
| | 470 | | 475 | | 480 |
| Ser Pro Ser Thr | Ala Ser Pro Tyr Cys | Arg Arg Thr Lys Ser | Asp | | |
| | 485 | | 490 | | 495 |
| Thr Glu Pro Gln | Lys Ser Gln Gln Ser | Ser Gly Arg Thr Ser | Gly | | |
| | 500 | | 505 | | 510 |
| Ser Asp Asp Pro | Gly Ile Cys Ser Asn | Thr Asp Ser Thr Gln | Ala | | |
| | 515 | | 520 | | 525 |
| Gln Val Leu Leu | Gly Lys Lys Arg Leu | Leu Lys Ala Glu Thr | Leu | | |
| | 530 | | 535 | | 540 |
| Glu Leu Ser Asp | Leu Tyr Val Ser Asp | Lys Lys Lys Asp Met | Ser | | |
| | 545 | | 550 | | 555 |
| Pro Pro Phe Ile | Cys Glu Glu Thr Asp | Glu Gln Lys Leu Gln | Thr | | |
| | 560 | | 565 | | 570 |
| Leu Asp Ile Gly | Ser Asn Leu Glu Lys | Glu Lys Leu Glu Asn | Ser | | |
| | 575 | | 580 | | 585 |
| Arg Ser Leu Glu | Cys Arg Ser Asp Pro | Glu Ser Pro Ile Lys | Lys | | |
| | 590 | | 595 | | 600 |
| Thr Ser Leu Ser | Pro Thr Ser Lys Leu | Gly Tyr Ser Tyr Ser | Arg | | |
| | 605 | | 610 | | 615 |
| Asp Leu Asp Leu | Ala Lys Lys Lys His | Ala Ser Leu Arg Gln | Thr | | |
| | 620 | | 625 | | 630 |
| Glu Ser Asp Pro | Asp Ala Asp Arg Thr | Thr Leu Asn His Ala | Asp | | |
| | 635 | | 640 | | 645 |
| His Ser Ser Lys | Ile Val Gln His Arg | Leu Leu Ser Arg Gln | Glu | | |
| | 650 | | 655 | | 660 |
| Glu Leu Lys Glu | Arg Ala Arg Val Leu | Leu Glu Gln Ala Arg | Arg | | |
| | 665 | | 670 | | 675 |
| Asp Ala Ala Leu | Lys Ala Gly Asn Lys | His Asn Thr Asn Thr | Ala | | |
| | 680 | | 685 | | 690 |
| Thr Pro Phe Cys | Asn Arg Gln Leu Ser | Asp Gln Gln Asp Glu | Glu | | |
| | 695 | | 700 | | 705 |
| Arg Arg Arg Gln | Leu Arg Glu Arg Ala | Arg Gln Leu Ile Ala | Glu | | |
| | 710 | | 715 | | 720 |
| Ala Arg Ser Gly | Val Lys Met Ser Glu | Leu Pro Ser Tyr Gly | Glu | | |
| | 725 | | 730 | | 735 |
| Met Ala Ala Glu | Lys Leu Lys Glu Arg | Ser Lys Ala Ser Gly | Asp | | |
| | 740 | | 745 | | 750 |
| Glu Asn Asp Asn | Ile Glu Ile Asp Thr | Asn Glu Glu Ile Pro | Glu | | |
| | 755 | | 760 | | 765 |
| Gly Phe Val Val | Gly Gly Gly Asp Glu | Leu Thr Asn Leu Glu | Asn | | |
| | 770 | | 775 | | 780 |
| Asp Leu Asp Thr | Pro Glu Gln Asn Ser | Lys Leu Val Asp Leu | Lys | | |
| | 785 | | 790 | | 795 |
| Leu Lys Lys Leu | Leu Glu Val Gln Pro | Gln Val Ala Asn Ser | Pro | | |
| | 800 | | 805 | | 810 |
| Ser Ser Ala Ala | Gln Lys Ala Val Thr | Glu Ser Ser Glu Gln | Asp | | |
| | 815 | | 820 | | 825 |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|------|-----|-----|-----|-----|------|-----|-----|-----|-----|------|--|
| Met | Lys | Ser | Gly | Thr | Glu | Asp | Leu | Arg | Thr | Glu | Arg | Leu | Gln | Lys | |
| | | | | 830 | | | | | 835 | | | | | 840 | |
| Thr | Thr | Glu | Arg | Phe | Arg | Asn | Pro | Val | Val | Phe | Ser | Lys | Asp | Ser | |
| | | | | 845 | | | | | 850 | | | | | 855 | |
| Thr | Val | Arg | Lys | Thr | Gln | Leu | Gln | Ser | Phe | Ser | Gln | Tyr | Ile | Glu | |
| | | | | 860 | | | | | 865 | | | | | 870 | |
| Asn | Arg | Pro | Glu | Met | Lys | Arg | Gln | Arg | Ser | Ile | Gln | Glu | Asp | Thr | |
| | | | | 875 | | | | | 880 | | | | | 885 | |
| Lys | Lys | Gly | Asn | Glu | Glu | Lys | Ala | Ala | Ile | Thr | Glu | Thr | Gln | Arg | |
| | | | | 890 | | | | | 895 | | | | | 900 | |
| Lys | Pro | Ser | Glu | Asp | Glu | Val | Leu | Asn | Lys | Gly | Phe | Lys | Asp | Thr | |
| | | | | 905 | | | | | 910 | | | | | 915 | |
| Ser | Gln | Tyr | Val | Val | Gly | Glu | Leu | Ala | Ala | Leu | Glu | Asn | Glu | Gln | |
| | | | | 920 | | | | | 925 | | | | | 930 | |
| Lys | Gln | Ile | Asp | Thr | Arg | Ala | Ala | Leu | Val | Glu | Lys | Arg | Leu | Arg | |
| | | | | 935 | | | | | 940 | | | | | 945 | |
| Tyr | Leu | Met | Asp | Thr | Gly | Arg | Asn | Thr | Glu | Glu | Glu | Glu | Ala | Met | |
| | | | | 950 | | | | | 955 | | | | | 960 | |
| Met | Gln | Glu | Trp | Phe | Met | Leu | Val | Asn | Lys | Lys | Asn | Ala | Leu | Ile | |
| | | | | 965 | | | | | 970 | | | | | 975 | |
| Arg | Arg | Met | Asn | Gln | Leu | Ser | Leu | Leu | Glu | Lys | Glu | His | Asp | Leu | |
| | | | | 980 | | | | | 985 | | | | | 990 | |
| Glu | Arg | Arg | Tyr | Glu | Leu | Leu | Asn | Arg | Glu | Leu | Arg | Ala | Met | Leu | |
| | | | | 995 | | | | | 1000 | | | | | 1005 | |
| Ala | Ile | Glu | Asp | Trp | Gln | Lys | Thr | Glu | Ala | Gln | Lys | Arg | Arg | Glu | |
| | | | | 1010 | | | | | 1015 | | | | | 1020 | |
| Gln | Leu | Leu | Leu | Asp | Glu | Leu | Val | Ala | Leu | Val | Asn | Lys | Arg | Asp | |
| | | | | 1025 | | | | | 1030 | | | | | 1035 | |
| Ala | Leu | Val | Arg | Asp | Leu | Asp | Ala | Gln | Glu | Lys | Gln | Ala | Glu | Glu | |
| | | | | 1040 | | | | | 1045 | | | | | 1050 | |
| Glu | Asp | Glu | His | Leu | Glu | Arg | Thr | Leu | Glu | Gln | Asn | Lys | Gly | Lys | |
| | | | | 1055 | | | | | 1060 | | | | | 1065 | |
| Met | Ala | Lys | Lys | Glu | Glu | Lys | Cys | Val | Leu | Gln | | | | | |
| | | | | 1070 | | | | | 1075 | | | | | | |

<210> 27

<211> 542

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5526375CD1

<400> 27

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Met | Glu | Glu | Arg | Gly | Ser | Pro | Asp | Gly | Asp | Leu | Ala | Arg | Ser | Leu | |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Glu | Gln | Gly | Pro | Glu | Gly | Pro | Glu | Thr | Pro | Ile | Gln | Val | Val | Leu | |
| | | | | 20 | | | | | 25 | | | | | 30 | |
| Arg | Val | Arg | Pro | Met | Ser | Ala | Ala | Glu | Leu | Arg | Arg | Gly | Gln | Gln | |
| | | | | 35 | | | | | 40 | | | | | 45 | |
| Ser | Val | Leu | His | Cys | Ser | Gly | Thr | Arg | Thr | Leu | Gln | Val | Ser | Pro | |
| | | | | 50 | | | | | 55 | | | | | 60 | |
| Pro | Gly | Gly | Gly | Pro | Glu | Val | Ala | Phe | Arg | Phe | Gly | Ala | Val | Leu | |
| | | | | 65 | | | | | 70 | | | | | 75 | |
| Asp | Ala | Ala | Arg | Thr | Gln | Glu | Asp | Val | Phe | Arg | Ala | Cys | Gly | Val | |
| | | | | 80 | | | | | 85 | | | | | 90 | |
| Arg | Arg | Leu | Gly | Glu | Leu | Ala | Leu | Arg | Gly | Phe | Ser | Cys | Thr | Val | |
| | | | | 95 | | | | | 100 | | | | | 105 | |
| Phe | Thr | Phe | Gly | Gln | Thr | Gly | Ser | Gly | Lys | Thr | Tyr | Thr | Leu | Thr | |
| | | | | 110 | | | | | 115 | | | | | 120 | |
| Gly | Pro | Pro | Pro | Gln | Gly | Glu | Gly | Val | Pro | Val | Pro | Pro | Ser | Leu | |
| | | | | 125 | | | | | 130 | | | | | 135 | |
| Ala | Gly | Ile | Met | Gln | Arg | Thr | Phe | Ala | Trp | Leu | Leu | Asp | Arg | Val | |
| | | | | 140 | | | | | 145 | | | | | 150 | |
| Gln | His | Leu | Gly | Ala | Pro | Val | Thr | Leu | Arg | Ala | Ser | Tyr | Leu | Glu | |
| | | | | 155 | | | | | 160 | | | | | 165 | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Ile | Tyr | Asn | Gly | Gln | Val | Arg | Asp | Leu | Leu | Ser | Leu | Gly | Ser | Pro | |
| | | | | 170 | | | | | 175 | | | | | 180 | |
| Arg | Pro | Leu | Pro | Val | Arg | Trp | Asn | Lys | Thr | Arg | Gly | Phe | Tyr | Val | |
| | | | | 185 | | | | | 190 | | | | | 195 | |
| Glu | Gln | Leu | Arg | Val | Val | Glu | Phe | Gly | Ser | Leu | Glu | Ala | Leu | Met | |
| | | | | 200 | | | | | 205 | | | | | 210 | |
| Glu | Leu | Leu | Gln | Thr | Gly | Leu | Ser | Arg | Arg | Arg | Asn | Ser | Ala | His | |
| | | | | 215 | | | | | 220 | | | | | 225 | |
| Thr | Leu | Asn | Gln | Ala | Ser | Ser | Arg | Ser | His | Ala | Leu | Leu | Thr | Leu | |
| | | | | 230 | | | | | 235 | | | | | 240 | |
| Tyr | Ile | Ser | Arg | Gln | Thr | Ala | Gln | Gln | Met | Pro | Ser | Val | Asp | Pro | |
| | | | | 245 | | | | | 250 | | | | | 255 | |
| Gly | Glu | Pro | Pro | Val | Gly | Gly | Lys | Leu | Cys | Phe | Val | Asp | Leu | Ala | |
| | | | | 260 | | | | | 265 | | | | | 270 | |
| Gly | Ser | Glu | Lys | Val | Ala | Ala | Thr | Gly | Ser | Arg | Gly | Glu | Leu | Met | |
| | | | | 275 | | | | | 280 | | | | | 285 | |
| Leu | Glu | Ala | Asn | Ser | Ile | Asn | Arg | Ser | Leu | Leu | Ala | Leu | Gly | His | |
| | | | | 290 | | | | | 295 | | | | | 300 | |
| Cys | Ile | Ser | Leu | Leu | Leu | Asp | Pro | Gln | Arg | Lys | Gln | Ser | His | Ile | |
| | | | | 305 | | | | | 310 | | | | | 315 | |
| Pro | Phe | Arg | Asp | Ser | Lys | Leu | Thr | Lys | Leu | Leu | Ala | Asp | Ser | Leu | |
| | | | | 320 | | | | | 325 | | | | | 330 | |
| Gly | Gly | Arg | Gly | Val | Thr | Leu | Met | Val | Ala | Cys | Val | Ser | Pro | Ser | |
| | | | | 335 | | | | | 340 | | | | | 345 | |
| Ala | Gln | Cys | Leu | Pro | Glu | Thr | Leu | Ser | Thr | Leu | Arg | Tyr | Ala | Ser | |
| | | | | 350 | | | | | 355 | | | | | 360 | |
| Arg | Ala | Gln | Arg | Val | Thr | Thr | Arg | Pro | Gln | Ala | Pro | Lys | Ser | Pro | |
| | | | | 365 | | | | | 370 | | | | | 375 | |
| Val | Ala | Lys | Gln | Pro | Gln | Arg | Leu | Glu | Thr | Glu | Met | Leu | Gln | Leu | |
| | | | | 380 | | | | | 385 | | | | | 390 | |
| Gln | Glu | Glu | Asn | Arg | Arg | Leu | Gln | Phe | Gln | Leu | Asp | Gln | Met | Asp | |
| | | | | 395 | | | | | 400 | | | | | 405 | |
| Cys | Lys | Ala | Ser | Gly | Leu | Ser | Gly | Ala | Arg | Val | Ala | Trp | Ala | Gln | |
| | | | | 410 | | | | | 415 | | | | | 420 | |
| Arg | Asn | Leu | Tyr | Gly | Met | Leu | Gln | Glu | Phe | Met | Leu | Glu | Asn | Glu | |
| | | | | 425 | | | | | 430 | | | | | 435 | |
| Arg | Leu | Arg | Lys | Glu | Lys | Ser | Gln | Leu | Gln | Asn | Ser | Arg | Glu | Leu | |
| | | | | 440 | | | | | 445 | | | | | 450 | |
| Ala | Gln | Asn | Glu | Gln | Arg | Ile | Leu | Ala | Gln | Gln | Val | His | Ala | Leu | |
| | | | | 455 | | | | | 460 | | | | | 465 | |
| Glu | Arg | Arg | Leu | Leu | Ser | Ala | Cys | Tyr | His | His | Gln | Gln | Gly | Pro | |
| | | | | 470 | | | | | 475 | | | | | 480 | |
| Gly | Leu | Thr | Pro | Pro | Cys | Pro | Cys | Leu | Met | Ala | Pro | Ala | Pro | Pro | |
| | | | | 485 | | | | | 490 | | | | | 495 | |
| Cys | His | Ala | Leu | Pro | Pro | Leu | Tyr | Ser | Cys | Pro | Cys | Cys | His | Ile | |
| | | | | 500 | | | | | 505 | | | | | 510 | |
| Cys | Pro | Leu | Cys | Arg | Val | Pro | Leu | Ala | His | Trp | Gly | Cys | Leu | Pro | |
| | | | | 515 | | | | | 520 | | | | | 525 | |
| Gly | Glu | His | His | Leu | Pro | Gln | Pro | Leu | Phe | Trp | Ala | Leu | Arg | Ser | |
| | | | | 530 | | | | | 535 | | | | | 540 | |

Gln Lys

<210> 28
 <211> 351
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5677408CD1

<400> 28
 Met Pro Ser Glu Thr Leu Trp Glu Ile Ala Lys Ala Glu Val Glu
 1 5 10 15
 Lys Arg Gly Ile Asn Gly Ser Glu Gly Asp Gly Ala Glu Ile Ala
 20 25 30

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Glu | Lys | Phe | Val | Phe | Phe | Ile | Gly | Ser | Lys | Asn | Gly | Gly | Lys | Thr | |
| | | | | 35 | | | | | 40 | | | | | | 45 |
| Thr | Ile | Ile | Leu | Arg | Cys | Leu | Asp | Arg | Asp | Glu | Pro | Pro | Lys | Pro | |
| | | | | 50 | | | | | 55 | | | | | | 60 |
| Thr | Leu | Ala | Leu | Glu | Tyr | Thr | Tyr | Gly | Arg | Arg | Ala | Lys | Gly | His | |
| | | | | 65 | | | | | 70 | | | | | | 75 |
| Asn | Thr | Pro | Lys | Asp | Ile | Ala | His | Phe | Trp | Glu | Leu | Gly | Gly | Gly | |
| | | | | 80 | | | | | 85 | | | | | | 90 |
| Thr | Ser | Leu | Leu | Asp | Leu | Ile | Ser | Ile | Pro | Ile | Thr | Gly | Asp | Thr | |
| | | | | 95 | | | | | 100 | | | | | | 105 |
| Leu | Arg | Thr | Phe | Ser | Leu | Val | Leu | Val | Leu | Asp | Leu | Ser | Lys | Pro | |
| | | | | 110 | | | | | 115 | | | | | | 120 |
| Asn | Asp | Leu | Trp | Pro | Thr | Met | Glu | Asn | Leu | Leu | Gln | Ala | Thr | Lys | |
| | | | | 125 | | | | | 130 | | | | | | 135 |
| Ser | His | Val | Asp | Lys | Val | Ile | Met | Lys | Leu | Gly | Lys | Thr | Asn | Ala | |
| | | | | 140 | | | | | 145 | | | | | | 150 |
| Lys | Ala | Val | Ser | Glu | Met | Arg | Gln | Lys | Ile | Trp | Asn | Asn | Met | Pro | |
| | | | | 155 | | | | | 160 | | | | | | 165 |
| Lys | Asp | His | Pro | Asp | His | Glu | Leu | Ile | Asp | Pro | Phe | Pro | Val | Pro | |
| | | | | 170 | | | | | 175 | | | | | | 180 |
| Leu | Val | Ile | Ile | Gly | Ser | Lys | Tyr | Asp | Val | Phe | Gln | Asp | Phe | Glu | |
| | | | | 185 | | | | | 190 | | | | | | 195 |
| Ser | Glu | Lys | Arg | Lys | Val | Ile | Cys | Lys | Thr | Leu | Arg | Phe | Val | Ala | |
| | | | | 200 | | | | | 205 | | | | | | 210 |
| His | Tyr | Tyr | Gly | Ala | Ser | Leu | Met | Phe | Thr | Ser | Lys | Ser | Glu | Ala | |
| | | | | 215 | | | | | 220 | | | | | | 225 |
| Leu | Leu | Leu | Lys | Ile | Arg | Gly | Val | Ile | Asn | Gln | Leu | Ala | Phe | Gly | |
| | | | | 230 | | | | | 235 | | | | | | 240 |
| Ile | Asp | Lys | Ser | Lys | Ser | Ile | Cys | Val | Asp | Gln | Asn | Lys | Pro | Leu | |
| | | | | 245 | | | | | 250 | | | | | | 255 |
| Phe | Ile | Thr | Ala | Gly | Leu | Asp | Ser | Phe | Gly | Gln | Ile | Gly | Ser | Pro | |
| | | | | 260 | | | | | 265 | | | | | | 270 |
| Pro | Val | Pro | Glu | Asn | Asp | Ile | Gly | Lys | Leu | His | Ala | His | Ser | Pro | |
| | | | | 275 | | | | | 280 | | | | | | 285 |
| Met | Glu | Leu | Trp | Lys | Lys | Val | Tyr | Glu | Lys | Leu | Phe | Pro | Pro | Lys | |
| | | | | 290 | | | | | 295 | | | | | | 300 |
| Ser | Ile | Asn | Thr | Leu | Lys | Asp | Ile | Lys | Asp | Pro | Ala | Arg | Asp | Pro | |
| | | | | 305 | | | | | 310 | | | | | | 315 |
| Gln | Tyr | Ala | Glu | Asn | Glu | Val | Asp | Glu | Met | Arg | Ile | Gln | Lys | Asp | |
| | | | | 320 | | | | | 325 | | | | | | 330 |
| Leu | Glu | Leu | Glu | Gln | Tyr | Lys | Arg | Ser | Ser | Ser | Lys | Ser | Trp | Lys | |
| | | | | 335 | | | | | 340 | | | | | | 345 |
| Gln | Ile | Glu | Leu | Asp | Ser | | | | | | | | | | |
| | | | | 350 | | | | | | | | | | | |

<210> 29

<211> 856

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5982278CD1

<400> 29

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Met | Lys | Ser | Ala | Arg | Ala | Lys | Thr | Pro | Arg | Lys | Pro | Thr | Val | Lys | |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Lys | Gly | Ser | Gln | Thr | Asn | Leu | Lys | Asp | Pro | Val | Gly | Val | Tyr | Cys | |
| | | | | 20 | | | | | 25 | | | | | 30 | |
| Arg | Val | Arg | Pro | Leu | Gly | Phe | Pro | Asp | Gln | Glu | Cys | Cys | Ile | Glu | |
| | | | | 35 | | | | | 40 | | | | | 45 | |
| Val | Ile | Asn | Asn | Thr | Thr | Val | Gln | Leu | His | Thr | Pro | Glu | Gly | Tyr | |
| | | | | 50 | | | | | 55 | | | | | 60 | |
| Arg | Leu | Asn | Arg | Asn | Gly | Asp | Tyr | Lys | Glu | Thr | Gln | Tyr | Ser | Phe | |
| | | | | 65 | | | | | 70 | | | | | 75 | |
| Lys | Gln | Val | Phe | Gly | Thr | His | Thr | Thr | Gln | Lys | Glu | Leu | Phe | Asp | |
| | | | | 80 | | | | | 85 | | | | | 90 | |

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Val | Val | Ala | Asn | Pro | Leu | Val | Asn | Asp | Leu | Ile | His | Gly | Lys | Asn |
| | | | | 95 | | | | | 100 | | | | | 105 |
| Gly | Leu | Leu | Phe | Thr | Tyr | Gly | Val | Thr | Gly | Ser | Gly | Lys | Thr | His |
| | | | | 110 | | | | | 115 | | | | | 120 |
| Thr | Met | Thr | Gly | Ser | Pro | Gly | Glu | Gly | Gly | Leu | Leu | Pro | Arg | Cys |
| | | | | 125 | | | | | 130 | | | | | 135 |
| Leu | Asp | Met | Ile | Phe | Asn | Ser | Ile | Gly | Ser | Phe | Gln | Ala | Lys | Arg |
| | | | | 140 | | | | | 145 | | | | | 150 |
| Tyr | Val | Phe | Lys | Ser | Asn | Asp | Arg | Asn | Ser | Met | Asp | Ile | Gln | Cys |
| | | | | 155 | | | | | 160 | | | | | 165 |
| Glu | Val | Asp | Ala | Leu | Leu | Glu | Arg | Gln | Lys | Arg | Glu | Ala | Met | Pro |
| | | | | 170 | | | | | 175 | | | | | 180 |
| Asn | Pro | Lys | Thr | Ser | Ser | Ser | Lys | Arg | Gln | Val | Asp | Pro | Glu | Phe |
| | | | | 185 | | | | | 190 | | | | | 195 |
| Ala | Asp | Met | Ile | Thr | Val | Gln | Glu | Phe | Cys | Lys | Ala | Glu | Glu | Val |
| | | | | 200 | | | | | 205 | | | | | 210 |
| Asp | Glu | Asp | Ser | Val | Tyr | Gly | Val | Phe | Val | Ser | Tyr | Ile | Glu | Ile |
| | | | | 215 | | | | | 220 | | | | | 225 |
| Tyr | Asn | Asn | Tyr | Ile | Tyr | Asp | Leu | Leu | Glu | Glu | Val | Pro | Phe | Asp |
| | | | | 230 | | | | | 235 | | | | | 240 |
| Pro | Ile | Asn | Pro | Asn | Leu | His | Asn | Leu | Asn | Cys | Phe | Val | Lys | Ile |
| | | | | 245 | | | | | 250 | | | | | 255 |
| Lys | Asn | His | Asn | Met | Tyr | Val | Ala | Gly | Cys | Thr | Glu | Val | Glu | Val |
| | | | | 260 | | | | | 265 | | | | | 270 |
| Lys | Ser | Thr | Glu | Glu | Ala | Phe | Glu | Val | Phe | Trp | Arg | Gly | Gln | Lys |
| | | | | 275 | | | | | 280 | | | | | 285 |
| Lys | Arg | Arg | Ile | Ala | Asn | Thr | His | Leu | Asn | Arg | Glu | Ser | Ser | Arg |
| | | | | 290 | | | | | 295 | | | | | 300 |
| Ser | His | Ser | Val | Phe | Asn | Ile | Lys | Leu | Val | Gln | Ala | Pro | Leu | Asp |
| | | | | 305 | | | | | 310 | | | | | 315 |
| Ala | Asp | Gly | Asp | Asn | Val | Leu | Gln | Glu | Lys | Glu | Gln | Ile | Thr | Ile |
| | | | | 320 | | | | | 325 | | | | | 330 |
| Ser | Gln | Leu | Ser | Leu | Val | Asp | Leu | Ala | Gly | Ser | Glu | Arg | Thr | Asn |
| | | | | 335 | | | | | 340 | | | | | 345 |
| Arg | Thr | Arg | Ala | Glu | Gly | Asn | Arg | Leu | Arg | Glu | Ala | Gly | Asn | Ile |
| | | | | 350 | | | | | 355 | | | | | 360 |
| Asn | Gln | Ser | Leu | Met | Thr | Leu | Arg | Thr | Cys | Met | Asp | Val | Leu | Arg |
| | | | | 365 | | | | | 370 | | | | | 375 |
| Glu | Asn | Gln | Met | Tyr | Gly | Thr | Asn | Lys | Met | Val | Pro | Tyr | Arg | Asp |
| | | | | 380 | | | | | 385 | | | | | 390 |
| Ser | Lys | Leu | Thr | His | Leu | Phe | Lys | Asn | Tyr | Phe | Asp | Gly | Glu | Gly |
| | | | | 395 | | | | | 400 | | | | | 405 |
| Lys | Val | Arg | Met | Ile | Val | Cys | Val | Asn | Pro | Lys | Ala | Glu | Asp | Tyr |
| | | | | 410 | | | | | 415 | | | | | 420 |
| Glu | Glu | Asn | Leu | Gln | Val | Met | Arg | Phe | Ala | Glu | Val | Thr | Gln | Glu |
| | | | | 425 | | | | | 430 | | | | | 435 |
| Val | Glu | Val | Ala | Arg | Pro | Val | Asp | Lys | Ala | Ile | Cys | Gly | Leu | Thr |
| | | | | 440 | | | | | 445 | | | | | 450 |
| Pro | Gly | Arg | Arg | Tyr | Arg | Asn | Gln | Pro | Arg | Gly | Pro | Val | Gly | Asn |
| | | | | 455 | | | | | 460 | | | | | 465 |
| Glu | Pro | Leu | Val | Thr | Asp | Val | Val | Leu | Gln | Ser | Phe | Pro | Pro | Leu |
| | | | | 470 | | | | | 475 | | | | | 480 |
| Pro | Ser | Cys | Glu | Ile | Leu | Asp | Ile | Asn | Asp | Glu | Gln | Thr | Leu | Pro |
| | | | | 485 | | | | | 490 | | | | | 495 |
| Arg | Leu | Ile | Glu | Ala | Leu | Glu | Lys | Arg | His | Asn | Leu | Arg | Gln | Met |
| | | | | 500 | | | | | 505 | | | | | 510 |
| Met | Ile | Asp | Glu | Phe | Asn | Lys | Gln | Ser | Asn | Ala | Phe | Lys | Ala | Leu |
| | | | | 515 | | | | | 520 | | | | | 525 |
| Leu | Gln | Glu | Phe | Asp | Asn | Ala | Val | Leu | Ser | Lys | Glu | Asn | His | Met |
| | | | | 530 | | | | | 535 | | | | | 540 |
| Gln | Gly | Lys | Leu | Asn | Glu | Lys | Glu | Lys | Met | Ile | Ser | Gly | Gln | Lys |
| | | | | 545 | | | | | 550 | | | | | 555 |
| Leu | Glu | Ile | Glu | Arg | Leu | Glu | Lys | Lys | Asn | Lys | Thr | Leu | Glu | Tyr |
| | | | | 560 | | | | | 565 | | | | | 570 |
| Lys | Ile | Glu | Ile | Leu | Glu | Lys | Thr | Thr | Thr | Ile | Tyr | Glu | Glu | Asp |
| | | | | 575 | | | | | 580 | | | | | 585 |
| Lys | Arg | Asn | Leu | Gln | Gln | Glu | Leu | Glu | Thr | Gln | Asn | Gln | Lys | Leu |

| | | | | | |
|-----------------|-----|---------------------|-----|---------------------|-----|
| Gln Arg Gln Phe | 590 | Ser Asp Lys Arg Arg | 595 | Leu Glu Ala Arg Leu | 600 |
| Gly Met Val Thr | 605 | Glu Thr Thr Met Lys | 610 | Glu Lys Glu Cys Glu | 615 |
| Arg Arg Val Ala | 620 | Ala Lys Gln Leu Glu | 625 | Gln Asn Lys Leu Trp | 630 |
| Val Lys Asp Glu | 635 | Lys Leu Lys Gln Leu | 640 | Lys Ala Ile Val Thr | 645 |
| Pro Lys Thr Glu | 650 | Lys Pro Glu Arg Pro | 655 | Ser Arg Glu Arg Asp | 660 |
| Glu Lys Val Thr | 665 | Gln Arg Ser Val Ser | 670 | Pro Ser Pro Val Pro | 675 |
| Leu Phe Gln Pro | 680 | Asp Gln Asn Ala Pro | 685 | Pro Ile Arg Leu Arg | 690 |
| Arg Arg Ser Arg | 695 | Ser Ala Gly Asp Arg | 700 | Val Asp His Lys | 705 |
| Ala Ser Asn Met | 710 | Gln Thr Glu Thr Val | 715 | Gln Pro His Val | 720 |
| His Ala Ile Thr | 725 | Val Ser Val Ala Asn | 730 | Glu Lys Ala Leu Ala | 735 |
| Cys Glu Lys Tyr | 740 | Met Leu Thr His Gln | 745 | Glu Leu Ala Ser Asp | 750 |
| Glu Ile Glu Thr | 755 | Lys Leu Ile Lys Gly | 760 | Asp Ile Tyr Lys Thr | 765 |
| Gly Gly Gly Gln | 770 | Ser Val Gln Phe Thr | 775 | Ile Glu Thr Leu Lys | 780 |
| Gln Glu Ser Pro | 785 | Asn Gly Ser Arg Lys | 790 | Arg Arg Ser Ser Thr | 795 |
| Ala Pro Ala Gln | 800 | Pro Asp Gly Ala Glu | 805 | Ser Glu Trp Thr Asp | 810 |
| Glu Thr Arg Cys | 815 | Ser Val Ala Val Glu | 820 | Met Arg Ala Gly Ser | 825 |
| Leu Gly Pro Gly | 830 | Tyr Gln His His Ala | 835 | Gln Pro Lys Arg Lys | 840 |
| Pro | 845 | | 850 | | 855 |

<210> 30

<211> 1056

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 6437362CD1

<400> 30

| | | | | | |
|-----------------|-----|-----------------|-----|-----------------|-----|
| Met Ala Cys Pro | Ala | Leu Gly Leu Glu | Ala | Leu Gln Pro Leu | Gln |
| 1 | 5 | 10 | 15 | | |
| Pro Glu Pro Pro | Pro | Glu Pro Ala Phe | Ser | Glu Ala Gln Lys | Trp |
| | 20 | 25 | 30 | | |
| Ile Glu Gln Val | Thr | Gly Arg Ser Phe | Gly | Asp Lys Asp Phe | Arg |
| | 35 | 40 | 45 | | |
| Thr Gly Leu Glu | Asn | Gly Ile Leu Leu | Cys | Glu Leu Leu Asn | Ala |
| | 50 | 55 | 60 | | |
| Ile Lys Pro Gly | Leu | Val Lys Lys Ile | Asn | Arg Leu Pro Thr | Pro |
| | 65 | 70 | 75 | | |
| Ile Ala Gly Leu | Asp | Asn Ile Ile Leu | Phe | Leu Arg Gly Cys | Lys |
| | 80 | 85 | 90 | | |
| Glu Leu Gly Leu | Lys | Glu Ser Gln Leu | Phe | Asp Pro Ser Asp | Leu |
| | 95 | 100 | 105 | | |
| Gln Asp Thr Ser | Asn | Arg Val Thr Val | Lys | Ser Leu Asp Tyr | Ser |
| | 110 | 115 | 120 | | |
| Arg Lys Leu Lys | Asn | Val Leu Val Thr | Ile | Tyr Trp Leu Gly | Lys |
| | 125 | 130 | 135 | | |
| Ala Ala Asn Ser | Cys | Thr Ser Tyr Ser | Gly | Thr Thr Leu Asn | Leu |

| | | | | | |
|---------------------|---------------------|---------------------|-----|--|-----|
| | 140 | | 145 | | 150 |
| Lys Glu Phe Glu Gly | Leu Leu Ala Gln Met | Arg Lys Asp Thr Asp | | | |
| | 155 | | 160 | | 165 |
| Asp Ile Glu Ser Pro | Lys Arg Ser Ile Arg | Asp Ser Gly Tyr Ile | | | |
| | 170 | | 175 | | 180 |
| Asp Cys Trp Asp Ser | Glu Arg Ser Asp Ser | Leu Ser Pro Pro Arg | | | |
| | 185 | | 190 | | 195 |
| His Gly Arg Asp Asp | Ser Phe Asp Ser Leu | Asp Ser Phe Gly Ser | | | |
| | 200 | | 205 | | 210 |
| Arg Ser Arg Gln Thr | Pro Ser Pro Asp Val | Val Leu Arg Gly Ser | | | |
| | 215 | | 220 | | 225 |
| Ser Asp Gly Arg Gly | Ser Asp Ser Glu Ser | Asp Leu Pro His Arg | | | |
| | 230 | | 235 | | 240 |
| Lys Leu Pro Asp Val | Lys Lys Asp Asp Met | Ser Ala Arg Arg Thr | | | |
| | 245 | | 250 | | 255 |
| Ser His Gly Glu Pro | Lys Ser Ala Val Pro | Phe Asn Gln Tyr Leu | | | |
| | 260 | | 265 | | 270 |
| Pro Asn Lys Ser Asn | Gln Thr Ala Tyr Val | Pro Ala Pro Leu Arg | | | |
| | 275 | | 280 | | 285 |
| Lys Lys Lys Ala Glu | Arg Glu Glu Tyr Arg | Lys Ser Trp Ser Thr | | | |
| | 290 | | 295 | | 300 |
| Ala Thr Ser Pro Leu | Gly Gly Glu Arg Pro | Phe Arg Tyr Gly Pro | | | |
| | 305 | | 310 | | 315 |
| Arg Thr Pro Val Ser | Asp Asp Ala Glu Ser | Thr Ser Met Phe Asp | | | |
| | 320 | | 325 | | 330 |
| Met Arg Cys Glu Glu | Glu Ala Ala Val Gln | Pro His Ser Arg Ala | | | |
| | 335 | | 340 | | 345 |
| Arg Gln Glu Gln Leu | Gln Leu Ile Asn Asn | Gln Leu Arg Glu Glu | | | |
| | 350 | | 355 | | 360 |
| Asp Asp Lys Trp Gln | Asp Asp Leu Ala Arg | Trp Lys Ser Arg Arg | | | |
| | 365 | | 370 | | 375 |
| Arg Ser Val Ser Gln | Asp Leu Ile Lys Lys | Glu Glu Glu Arg Lys | | | |
| | 380 | | 385 | | 390 |
| Lys Met Glu Lys Leu | Leu Ala Gly Glu Asp | Gly Thr Ser Glu Arg | | | |
| | 395 | | 400 | | 405 |
| Arg Lys Ser Ile Lys | Thr Tyr Arg Glu Ile | Val Gln Glu Lys Glu | | | |
| | 410 | | 415 | | 420 |
| Arg Arg Glu Arg Glu | Leu His Glu Ala Tyr | Lys Asn Ala Arg Ser | | | |
| | 425 | | 430 | | 435 |
| Gln Glu Glu Ala Glu | Gly Ile Leu Gln Gln | Tyr Ile Glu Arg Phe | | | |
| | 440 | | 445 | | 450 |
| Thr Ile Ser Glu Ala | Val Leu Glu Arg Leu | Glu Met Pro Lys Ile | | | |
| | 455 | | 460 | | 465 |
| Leu Glu Arg Ser His | Ser Thr Glu Pro Asn | Leu Ser Ser Phe Leu | | | |
| | 470 | | 475 | | 480 |
| Asn Asp Pro Asn Pro | Met Lys Tyr Leu Arg | Gln Gln Ser Leu Pro | | | |
| | 485 | | 490 | | 495 |
| Pro Pro Lys Phe Thr | Ala Thr Val Glu Thr | Thr Ile Ala Arg Ala | | | |
| | 500 | | 505 | | 510 |
| Ser Val Leu Asp Thr | Ser Met Ser Ala Gly | Ser Gly Ser Pro Ser | | | |
| | 515 | | 520 | | 525 |
| Lys Thr Val Thr Pro | Lys Ala Val Pro Met | Leu Thr Pro Lys Pro | | | |
| | 530 | | 535 | | 540 |
| Tyr Ser Gln Pro Lys | Asn Ser Gln Asp Val | Leu Lys Thr Phe Lys | | | |
| | 545 | | 550 | | 555 |
| Val Asp Gly Lys Val | Ser Val Asn Gly Glu | Thr Val His Arg Glu | | | |
| | 560 | | 565 | | 570 |
| Glu Glu Lys Glu Arg | Glu Cys Pro Thr Val | Ala Pro Ala His Ser | | | |
| | 575 | | 580 | | 585 |
| Leu Thr Lys Ser Gln | Met Phe Glu Gly Val | Ala Arg Val His Gly | | | |
| | 590 | | 595 | | 600 |
| Ser Pro Leu Glu Leu | Lys Gln Asp Asn Gly | Ser Ile Glu Ile Asn | | | |
| | 605 | | 610 | | 615 |
| Ile Lys Lys Pro Asn | Ser Val Pro Gln Glu | Leu Ala Ala Thr Thr | | | |
| | 620 | | 625 | | 630 |
| Glu Lys Thr Glu Pro | Asn Ser Gln Glu Asp | Lys Asn Asp Gly Gly | | | |
| | 635 | | 640 | | 645 |

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|------|-----|-----|-----|-----|------|-----|-----|-----|-----|------|
| Lys | Ser | Arg | Lys | Gly | Asn | Ile | Glu | Leu | Ala | Ser | Ser | Glu | Pro | Gln |
| | | | | 650 | | | | | 655 | | | | | 660 |
| His | Phe | Thr | Thr | Thr | Val | Thr | Arg | Cys | Ser | Pro | Thr | Val | Ala | Phe |
| | | | | 665 | | | | | 670 | | | | | 675 |
| Val | Glu | Phe | Pro | Ser | Ser | Pro | Gln | Leu | Lys | Asn | Asp | Val | Ser | Glu |
| | | | | 680 | | | | | 685 | | | | | 690 |
| Glu | Lys | Asp | Gln | Lys | Lys | Pro | Glu | Asn | Glu | Met | Ser | Gly | Lys | Val |
| | | | | 695 | | | | | 700 | | | | | 705 |
| Glu | Leu | Val | Leu | Ser | Gln | Lys | Val | Val | Lys | Pro | Lys | Ser | Pro | Glu |
| | | | | 710 | | | | | 715 | | | | | 720 |
| Pro | Glu | Ala | Thr | Leu | Thr | Phe | Pro | Phe | Leu | Asp | Lys | Met | Pro | Glu |
| | | | | 725 | | | | | 730 | | | | | 735 |
| Ala | Asn | Gln | Leu | His | Leu | Pro | Asn | Leu | Asn | Ser | Gln | Val | Asp | Ser |
| | | | | 740 | | | | | 745 | | | | | 750 |
| Pro | Ser | Ser | Glu | Lys | Ser | Pro | Val | Thr | Thr | Pro | Phe | Lys | Phe | Trp |
| | | | | 755 | | | | | 760 | | | | | 765 |
| Ala | Trp | Asp | Pro | Glu | Glu | Glu | Arg | Arg | Arg | Gln | Glu | Lys | Trp | Gln |
| | | | | 770 | | | | | 775 | | | | | 780 |
| Gln | Glu | Gln | Glu | Arg | Leu | Leu | Gln | Glu | Arg | Tyr | Gln | Lys | Glu | Gln |
| | | | | 785 | | | | | 790 | | | | | 795 |
| Asp | Lys | Leu | Lys | Glu | Glu | Trp | Glu | Lys | Ala | Gln | Lys | Glu | Val | Glu |
| | | | | 800 | | | | | 805 | | | | | 810 |
| Glu | Glu | Glu | Arg | Arg | Tyr | Tyr | Glu | Glu | Glu | Arg | Lys | Ile | Ile | Glu |
| | | | | 815 | | | | | 820 | | | | | 825 |
| Asp | Thr | Val | Val | Pro | Phe | Thr | Val | Ser | Ser | Ser | Ser | Ala | Asp | Gln |
| | | | | 830 | | | | | 835 | | | | | 840 |
| Leu | Ser | Thr | Ser | Ser | Ser | Met | Thr | Glu | Gly | Ser | Gly | Thr | Met | Asn |
| | | | | 845 | | | | | 850 | | | | | 855 |
| Lys | Ile | Asp | Leu | Gly | Asn | Cys | Gln | Asp | Glu | Lys | Gln | Asp | Arg | Arg |
| | | | | 860 | | | | | 865 | | | | | 870 |
| Trp | Lys | Lys | Ser | Phe | Gln | Gly | Asp | Asp | Ser | Asp | Leu | Leu | Leu | Lys |
| | | | | 875 | | | | | 880 | | | | | 885 |
| Thr | Arg | Glu | Ser | Asp | Arg | Leu | Glu | Glu | Lys | Gly | Ser | Leu | Thr | Glu |
| | | | | 890 | | | | | 895 | | | | | 900 |
| Gly | Ala | Leu | Ala | His | Ser | Gly | Asn | Pro | Val | Ser | Lys | Gly | Val | His |
| | | | | 905 | | | | | 910 | | | | | 915 |
| Glu | Asp | His | Gln | Leu | Asp | Thr | Glu | Ala | Gly | Ala | Pro | His | Cys | Gly |
| | | | | 920 | | | | | 925 | | | | | 930 |
| Thr | Asn | Pro | Gln | Leu | Ala | Gln | Asp | Pro | Ser | Gln | Asn | Gln | Gln | Thr |
| | | | | 935 | | | | | 940 | | | | | 945 |
| Ser | Asn | Pro | Thr | His | Ser | Ser | Glu | Asp | Val | Lys | Pro | Lys | Thr | Leu |
| | | | | 950 | | | | | 955 | | | | | 960 |
| Pro | Leu | Asp | Lys | Ser | Ile | Asn | His | Gln | Ile | Glu | Ser | Pro | Ser | Glu |
| | | | | 965 | | | | | 970 | | | | | 975 |
| Arg | Arg | Lys | Ser | Ile | Ser | Gly | Lys | Lys | Leu | Cys | Ser | Ser | Cys | Gly |
| | | | | 980 | | | | | 985 | | | | | 990 |
| Leu | Pro | Leu | Gly | Lys | Gly | Ala | Ala | Met | Ile | Ile | Glu | Thr | Leu | Asn |
| | | | | 995 | | | | | 1000 | | | | | 1005 |
| Leu | Tyr | Phe | His | Ile | Gln | Cys | Phe | Arg | Cys | Gly | Ile | Cys | Lys | Gly |
| | | | | 1010 | | | | | 1015 | | | | | 1020 |
| Gln | Leu | Gly | Asp | Ala | Val | Ser | Gly | Thr | Asp | Val | Arg | Ile | Arg | Asn |
| | | | | 1025 | | | | | 1030 | | | | | 1035 |
| Gly | Leu | Leu | Asn | Cys | Asn | Asp | Cys | Tyr | Met | Arg | Ser | Arg | Ser | Ala |
| | | | | 1040 | | | | | 1045 | | | | | 1050 |
| Gly | Gln | Pro | Thr | Thr | Leu | | | | | | | | | |
| | | | | 1055 | | | | | | | | | | |

<210> 31

<211> 1569

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4173970CD1

<400> 31

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Val | Val | Pro | Pro | Gln | Glu | Pro | Asp | Arg | Thr | Ser | Gln | Glu | Asn |
| 1 | | | | 5 | | | | | 10 | | | | | 15 |
| Ser | Pro | Ala | Leu | Leu | Gly | Val | Gln | Lys | Ala | Val | Ser | Thr | Arg | Val |
| | | | | 20 | | | | | 25 | | | | | 30 |
| Pro | Thr | Gly | Ser | Asn | Ser | Ser | Ser | Gln | Thr | Thr | Glu | Cys | Leu | Thr |
| | | | | 35 | | | | | 40 | | | | | 45 |
| Pro | Glu | Ser | Cys | Ser | Gln | Thr | Thr | Ser | Asn | Val | Ala | Ser | Gln | Ser |
| | | | | 50 | | | | | 55 | | | | | 60 |
| Met | Pro | Pro | Val | Tyr | Pro | Ser | Val | Asp | Ile | Asp | Ala | His | Thr | Glu |
| | | | | 65 | | | | | 70 | | | | | 75 |
| Ser | Asn | His | Asp | Thr | Ala | Leu | Thr | Leu | Ala | Cys | Ala | Gly | Gly | His |
| | | | | 80 | | | | | 85 | | | | | 90 |
| Glu | Glu | Leu | Val | Ser | Val | Leu | Ile | Ala | Arg | Asp | Ala | Lys | Ile | Glu |
| | | | | 95 | | | | | 100 | | | | | 105 |
| His | Arg | Asp | Lys | Lys | Gly | Phe | Thr | Pro | Leu | Ile | Leu | Ala | Ala | Thr |
| | | | | 110 | | | | | 115 | | | | | 120 |
| Ala | Gly | His | Val | Gly | Val | Val | Glu | Ile | Leu | Leu | Asp | Lys | Gly | Gly |
| | | | | 125 | | | | | 130 | | | | | 135 |
| Asp | Ile | Glu | Ala | Gln | Ser | Glu | Arg | Thr | Lys | Asp | Thr | Pro | Leu | Ser |
| | | | | 140 | | | | | 145 | | | | | 150 |
| Leu | Ala | Cys | Ser | Gly | Gly | Arg | Gln | Glu | Val | Val | Asp | Leu | Leu | Leu |
| | | | | 155 | | | | | 160 | | | | | 165 |
| Ala | Arg | Gly | Ala | Asn | Lys | Glu | His | Arg | Asn | Val | Ser | Asp | Tyr | Thr |
| | | | | 170 | | | | | 175 | | | | | 180 |
| Pro | Leu | Ser | Leu | Ala | Ala | Ser | Gly | Gly | Tyr | Val | Asn | Ile | Ile | Lys |
| | | | | 185 | | | | | 190 | | | | | 195 |
| Ile | Leu | Leu | Asn | Ala | Gly | Ala | Glu | Ile | Asn | Ser | Arg | Thr | Gly | Ser |
| | | | | 200 | | | | | 205 | | | | | 210 |
| Lys | Leu | Gly | Ile | Ser | Pro | Leu | Met | Leu | Ala | Ala | Met | Asn | Gly | His |
| | | | | 215 | | | | | 220 | | | | | 225 |
| Val | Pro | Ala | Val | Lys | Leu | Leu | Leu | Asp | Met | Gly | Ser | Asp | Ile | Asn |
| | | | | 230 | | | | | 235 | | | | | 240 |
| Ala | Gln | Ile | Glu | Thr | Asn | Arg | Asn | Thr | Ala | Leu | Thr | Leu | Ala | Cys |
| | | | | 245 | | | | | 250 | | | | | 255 |
| Phe | Gln | Gly | Arg | Ala | Glu | Val | Val | Ser | Leu | Leu | Leu | Asp | Arg | Lys |
| | | | | 260 | | | | | 265 | | | | | 270 |
| Ala | Asn | Val | Glu | His | Arg | Ala | Lys | Thr | Gly | Leu | Thr | Pro | Leu | Met |
| | | | | 275 | | | | | 280 | | | | | 285 |
| Glu | Ala | Ala | Ser | Gly | Gly | Tyr | Ala | Glu | Val | Gly | Arg | Val | Leu | Leu |
| | | | | 290 | | | | | 295 | | | | | 300 |
| Asp | Lys | Gly | Ala | Asp | Val | Asn | Ala | Pro | Pro | Val | Pro | Ser | Ser | Arg |
| | | | | 305 | | | | | 310 | | | | | 315 |
| Asp | Thr | Ala | Leu | Thr | Ile | Ala | Ala | Asp | Lys | Gly | His | Tyr | Lys | Phe |
| | | | | 320 | | | | | 325 | | | | | 330 |
| Cys | Glu | Leu | Leu | Ile | His | Arg | Gly | Ala | His | Ile | Asp | Val | Arg | Asn |
| | | | | 335 | | | | | 340 | | | | | 345 |
| Lys | Lys | Gly | Asn | Thr | Pro | Leu | Trp | Leu | Ala | Ser | Asn | Gly | Gly | His |
| | | | | 350 | | | | | 355 | | | | | 360 |
| Phe | Asp | Val | Val | Gln | Leu | Leu | Val | Gln | Ala | Gly | Ala | Asp | Val | Asp |
| | | | | 365 | | | | | 370 | | | | | 375 |
| Ala | Ala | Asp | Asn | Arg | Lys | Ile | Thr | Pro | Leu | Met | Ser | Ala | Phe | Arg |
| | | | | 380 | | | | | 385 | | | | | 390 |
| Lys | Gly | His | Val | Lys | Val | Val | Gln | Tyr | Leu | Val | Lys | Glu | Val | Asn |
| | | | | 395 | | | | | 400 | | | | | 405 |
| Gln | Phe | Pro | Ser | Asp | Ile | Glu | Cys | Met | Arg | Tyr | Ile | Ala | Thr | Ile |
| | | | | 410 | | | | | 415 | | | | | 420 |
| Thr | Asp | Lys | Glu | Leu | Leu | Lys | Lys | Cys | His | Gln | Cys | Val | Glu | Thr |
| | | | | 425 | | | | | 430 | | | | | 435 |
| Ile | Val | Lys | Ala | Lys | Asp | Gln | Gln | Ala | Ala | Glu | Ala | Asn | Lys | Asn |
| | | | | 440 | | | | | 445 | | | | | 450 |
| Ala | Ser | Ile | Leu | Leu | Lys | Glu | Leu | Asp | Leu | Glu | Lys | Ser | Arg | Glu |
| | | | | 455 | | | | | 460 | | | | | 465 |
| Glu | Ser | Arg | Lys | Gln | Ala | Leu | Ala | Ala | Lys | Arg | Glu | Lys | Arg | Lys |
| | | | | 470 | | | | | 475 | | | | | 480 |
| Glu | Lys | Arg | Lys | Lys | Lys | Lys | Glu | Glu | Gln | Lys | Arg | Lys | Gln | Glu |
| | | | | 485 | | | | | 490 | | | | | 495 |
| Glu | Asp | Glu | Glu | Asn | Lys | Pro | Lys | Glu | Asn | Ser | Glu | Leu | Pro | Glu |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|------|
| Asp | Glu | Asp | Glu | 500 | Glu | Glu | Asn | Asp | Glu | 505 | Val | Glu | Gln | Glu | 510 |
| Pro | Ile | Glu | Pro | 515 | Pro | Ser | Ala | Thr | Thr | 520 | Thr | Thr | Ile | Gly | 525 |
| Ser | Ala | Thr | Ser | 530 | Ala | Thr | Phe | Thr | Asn | 535 | Val | Phe | Gly | Lys | 540 |
| Ala | Asn | Val | Val | 545 | Thr | Thr | Pro | Ser | Thr | 550 | Asn | Arg | Lys | Asn | 555 |
| Asn | Lys | Thr | Lys | 560 | Glu | Thr | Pro | Pro | Thr | 565 | Ala | His | Leu | Ile | 570 |
| Glu | Gln | His | Met | 575 | Ser | Leu | Ala | Gln | Gln | 580 | Lys | Ala | Asp | Lys | 585 |
| Ile | Asn | Gly | Glu | 590 | Pro | Arg | Gly | Gly | Gly | 595 | Ala | Gly | Gly | Asn | 600 |
| Ser | Asp | Asn | Leu | 605 | Asp | Ser | Thr | Asp | Cys | 610 | Asn | Ser | Glu | Ser | 615 |
| Gly | Gly | Lys | Ser | 620 | Gln | Glu | Leu | Asn | Phe | 625 | Val | Met | Asp | Val | 630 |
| Ser | Lys | Tyr | Pro | 635 | Ser | Leu | Leu | Leu | His | 640 | Ser | Gln | Glu | Glu | 645 |
| Ser | Thr | Ala | Thr | 650 | Ser | Lys | Thr | Gln | Thr | 655 | Arg | Tyr | Lys | Thr | 660 |
| Leu | Pro | Leu | Ser | 665 | Ser | Pro | Asn | Ile | Lys | 670 | Leu | Asn | Leu | Thr | 675 |
| Lys | Arg | Gly | Gln | 680 | Lys | Arg | Glu | Glu | Gly | 685 | Trp | Lys | Glu | Val | 690 |
| Arg | Ser | Lys | Lys | 695 | Leu | Ser | Val | Pro | Ala | 700 | Ser | Val | Val | Ser | 705 |
| Met | Gly | Arg | Gly | 710 | Gly | Cys | Asn | Ile | Thr | 715 | Ala | Ile | Gln | Asp | 720 |
| Gly | Ala | His | Ile | 725 | Asp | Val | Asp | Lys | Gln | 730 | Lys | Asp | Lys | Asn | 735 |
| Arg | Met | Ile | Thr | 740 | Ile | Arg | Gly | Gly | Thr | 745 | Glu | Ser | Thr | Arg | 750 |
| Val | Gln | Leu | Ile | 755 | Asn | Ala | Leu | Ile | Gln | 760 | Asp | Pro | Ala | Lys | 765 |
| Glu | Asp | Leu | Ile | 770 | Pro | Lys | Asn | His | Ile | 775 | Arg | Thr | Pro | Ala | 780 |
| Lys | Ser | Ile | His | 785 | Ala | Asn | Phe | Ser | Ser | 790 | Gly | Val | Gly | Thr | 795 |
| Ala | Ser | Ser | Lys | 800 | Asn | Ala | Phe | Pro | Leu | 805 | Gly | Ala | Pro | Thr | 810 |
| Thr | Ser | Gln | Ala | 815 | Thr | Thr | Leu | Ser | Thr | 820 | Phe | Gln | Pro | Ala | 825 |
| Leu | Asn | Lys | Asn | 830 | Val | Pro | Thr | Asn | Val | 835 | Arg | Ser | Ser | Phe | 840 |
| Ser | Leu | Pro | Leu | 845 | Ala | Tyr | Pro | His | Pro | 850 | His | Phe | Ala | Leu | 855 |
| Ala | Gln | Thr | Met | 860 | Gln | Gln | Ile | Arg | His | 865 | Pro | Arg | Leu | Pro | 870 |
| Gln | Phe | Gly | Gly | 875 | Thr | Phe | Ser | Pro | Ser | 880 | Pro | Asn | Thr | Trp | 885 |
| Phe | Pro | Val | Arg | 890 | Pro | Val | Asn | Pro | Gly | 895 | Asn | Thr | Asn | Ser | 900 |
| Lys | His | Asn | Asn | 905 | Thr | Ser | Arg | Leu | Pro | 910 | Asn | Gln | Asn | Gly | 915 |
| Leu | Pro | Ser | Glu | 920 | Ser | Ala | Gly | Leu | Ala | 925 | Thr | Ala | Ser | Cys | 930 |
| Thr | Val | Ser | Ser | 935 | Val | Val | Ala | Ala | Ser | 940 | Gln | Gln | Leu | Cys | 945 |
| Asn | Thr | Arg | Thr | 950 | Pro | Ser | Ser | Val | Arg | 955 | Lys | Gln | Leu | Phe | 960 |
| Val | Pro | Lys | Thr | 965 | Ser | Pro | Pro | Ala | Thr | 970 | Val | Ile | Ser | Ser | 975 |
| Ser | Thr | Cys | Ser | 980 | Ser | Leu | Pro | Ser | Val | 985 | Ser | Ser | Ala | Pro | 990 |
| | | | | 995 | | | | | | 1000 | | | | | 1005 |

| | | |
|---------------------|---------------------|---------------------|
| Ser Gly Gln Ala Pro | Thr Thr Phe Leu Pro | Ala Ser Thr Ser Gln |
| 1010 | 1015 | 1020 |
| Ala Gln Leu Ser Ser | Gln Lys Met Glu Ser | Phe Ser Ala Val Pro |
| 1025 | 1030 | 1035 |
| Pro Thr Lys Glu Lys | Val Ser Thr Gln Asp | Gln Pro Met Ala Asn |
| 1040 | 1045 | 1050 |
| Leu Cys Thr Pro Ser | Ser Thr Ala Asn Ser | Cys Ser Ser Ser Ala |
| 1055 | 1060 | 1065 |
| Ser Asn Thr Pro Gly | Ala Pro Glu Thr His | Pro Ser Ser Ser Pro |
| 1070 | 1075 | 1080 |
| Thr Pro Thr Ser Ser | Asn Thr Gln Glu Glu | Ala Gln Pro Ser Ser |
| 1085 | 1090 | 1095 |
| Val Ser Asp Leu Ser | Pro Met Ser Met Pro | Phe Ala Ser Asn Ser |
| 1100 | 1105 | 1110 |
| Glu Pro Ala Pro Leu | Thr Leu Thr Ser Pro | Arg Met Val Ala Ala |
| 1115 | 1120 | 1125 |
| Asp Asn Gln Asp Thr | Ser Asn Leu Pro Gln | Leu Ala Val Pro Ala |
| 1130 | 1135 | 1140 |
| Pro Arg Val Ser His | Arg Met Gln Pro Arg | Gly Ser Phe Tyr Ser |
| 1145 | 1150 | 1155 |
| Met Val Pro Asn Ala | Thr Ile His Gln Asp | Pro Gln Ser Ile Phe |
| 1160 | 1165 | 1170 |
| Val Thr Asn Pro Val | Thr Leu Thr Pro Pro | Gln Gly Pro Pro Ala |
| 1175 | 1180 | 1185 |
| Ala Val Gln Leu Ser | Ser Ala Val Asn Ile | Met Asn Gly Ser Gln |
| 1190 | 1195 | 1200 |
| Met His Ile Asn Pro | Ala Asn Lys Ser Leu | Pro Pro Thr Phe Gly |
| 1205 | 1210 | 1215 |
| Pro Ala Thr Leu Phe | Asn His Phe Ser Ser | Leu Phe Asp Ser Ser |
| 1220 | 1225 | 1230 |
| Gln Val Pro Ala Asn | Gln Gly Trp Gly Asp | Gly Pro Leu Ser Ser |
| 1235 | 1240 | 1245 |
| Arg Val Ala Thr Asp | Ala Ser Phe Thr Val | Gln Ser Ala Phe Leu |
| 1250 | 1255 | 1260 |
| Gly Asn Ser Val Leu | Gly His Leu Glu Asn | Met His Pro Asp Asn |
| 1265 | 1270 | 1275 |
| Ser Lys Ala Pro Gly | Phe Arg Pro Pro Ser | Gln Arg Val Ser Thr |
| 1280 | 1285 | 1290 |
| Ser Pro Val Gly Leu | Pro Ser Ile Asp Pro | Ser Gly Ser Ser Pro |
| 1295 | 1300 | 1305 |
| Ser Ser Ser Ser Ala | Pro Leu Ala Ser Phe | Ser Gly Ile Pro Gly |
| 1310 | 1315 | 1320 |
| Thr Arg Val Phe Leu | Gln Gly Pro Ala Pro | Val Gly Thr Pro Ser |
| 1325 | 1330 | 1335 |
| Phe Asn Arg Gln His | Phe Ser Pro His Pro | Trp Thr Ser Ala Ser |
| 1340 | 1345 | 1350 |
| Asn Ser Cys Asp Ser | Pro Ile Pro Ser Val | Ser Ser Gly Ser Ser |
| 1355 | 1360 | 1365 |
| Ser Pro Leu Ser Ala | Thr Ser Ala Pro Pro | Thr Leu Gly Gln Pro |
| 1370 | 1375 | 1380 |
| Lys Gly Val Ser Ala | Ser Gln Asp Arg Lys | Ile Pro Pro Pro Ile |
| 1385 | 1390 | 1395 |
| Gly Thr Glu Arg Leu | Ala Arg Ile Arg Gln | Gly Gly Ser Val Ala |
| 1400 | 1405 | 1410 |
| Gln Ala Pro Ala Gly | Thr Ser Phe Val Ala | Pro Val Gly His Ser |
| 1415 | 1420 | 1425 |
| Gly Ile Trp Ser Phe | Gly Val Asn Ala Val | Ser Glu Gly Leu Ser |
| 1430 | 1435 | 1440 |
| Gly Trp Ser Gln Ser | Val Met Gly Asn His | Pro Met His Gln Gln |
| 1445 | 1450 | 1455 |
| Leu Ser Asp Pro Ser | Thr Phe Ser Gln His | Gln Pro Met Glu Arg |
| 1460 | 1465 | 1470 |
| Asp Asp Ser Gly Met | Val Ala Pro Ser Asn | Ile Phe His Gln Pro |
| 1475 | 1480 | 1485 |
| Met Ala Ser Gly Phe | Val Asp Phe Ser Lys | Gly Leu Pro Ile Ser |
| 1490 | 1495 | 1500 |
| Met Tyr Gly Gly Thr | Ile Ile Pro Ser His | Pro Gln Leu Ala Asp |

| | | |
|---------------------|---------------------|---------------------|
| 1505 | 1510 | 1515 |
| Val Pro Gly Gly Pro | Leu Phe Asn Gly Leu | His Asn Pro Asp Pro |
| 1520 | 1525 | 1530 |
| Ala Trp Asn Pro Met | Ile Lys Val Ile Gln | Asn Ser Thr Glu Cys |
| 1535 | 1540 | 1545 |
| Thr Asp Ala Gln Gln | Ile Trp Pro Gly Thr | Trp Ala Pro His Ile |
| 1550 | 1555 | 1560 |
| Gly Asn Met His Leu | Lys Tyr Val Asn | |
| 1565 | | |

<210> 32
 <211> 680
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2772751CD1

<400> 32

| | | |
|---------------------|---------------------|---------------------|
| Met Val Phe Leu Leu | Leu Thr Thr Leu Leu | Leu Leu Ile Gly Val |
| 1 | 5 | 10 |
| Val Cys Ala Phe Val | Thr Asn Gln Arg Thr | His Glu Gln Met Gly |
| 20 | 25 | 30 |
| Pro Ser Ile Glu Ala | Met Pro Glu Thr Leu | Leu Ser Leu Trp Gly |
| 35 | 40 | 45 |
| Leu Val Ser Asp Val | Pro Gln Glu Leu Gln | Ala Val Ala Gln Gln |
| 50 | 55 | 60 |
| Phe Ser Leu Pro Gln | Glu Gln Val Ser Glu | Glu Leu Asp Gly Val |
| 65 | 70 | 75 |
| Gly Val Ser Ile Gly | Ser Ala Ile His Thr | Gln Leu Arg Ser Ser |
| 80 | 85 | 90 |
| Val Tyr Pro Leu Leu | Ala Ala Val Gly Ser | Leu Gly Gln Val Leu |
| 95 | 100 | 105 |
| Gln Val Ser Val His | His Leu Gln Thr Leu | Asn Ala Thr Val Val |
| 110 | 115 | 120 |
| Glu Leu Gln Ala Gly | Gln Gln Asp Leu Glu | Pro Ala Ile Arg Glu |
| 125 | 130 | 135 |
| His Arg Asp Arg Leu | Leu Glu Leu Leu Gln | Glu Ala Arg Cys Gln |
| 140 | 145 | 150 |
| Gly Asp Cys Ala Gly | Ala Leu Ser Trp Ala | Arg Thr Leu Glu Leu |
| 155 | 160 | 165 |
| Gly Ala Asp Phe Ser | Gln Val Pro Ser Val | Asp His Val Leu His |
| 170 | 175 | 180 |
| Gln Leu Lys Gly Val | Pro Glu Ala Asn Phe | Ser Ser Met Val Gln |
| 185 | 190 | 195 |
| Glu Glu Asn Ser Thr | Phe Asn Ala Leu Pro | Ala Leu Ala Ala Met |
| 200 | 205 | 210 |
| Gln Thr Ser Ser Val | Val Gln Glu Leu Lys | Lys Ala Val Ala Gln |
| 215 | 220 | 225 |
| Gln Pro Glu Gly Val | Arg Thr Leu Ala Glu | Gly Phe Pro Gly Leu |
| 230 | 235 | 240 |
| Glu Ala Ala Ser Arg | Trp Ala Gln Ala Leu | Gln Glu Val Glu Glu |
| 245 | 250 | 255 |
| Ser Ser Arg Pro Tyr | Leu Gln Glu Val Gln | Arg Tyr Glu Thr Tyr |
| 260 | 265 | 270 |
| Arg Trp Ile Val Gly | Cys Val Leu Cys Ser | Val Val Leu Phe Val |
| 275 | 280 | 285 |
| Val Leu Cys Asn Leu | Leu Gly Leu Asn Leu | Gly Ile Trp Gly Leu |
| 290 | 295 | 300 |
| Ser Ala Arg Asp Asp | Pro Ser His Pro Glu | Ala Lys Gly Glu Ala |
| 305 | 310 | 315 |
| Gly Ala Arg Phe Leu | Met Ala Gly Val Gly | Leu Ser Phe Leu Phe |
| 320 | 325 | 330 |
| Ala Ala Pro Leu Ile | Leu Leu Val Phe Ala | Thr Phe Leu Val Gly |
| 335 | 340 | 345 |
| Gly Asn Val Gln Thr | Leu Val Cys Gln Ser | Trp Glu Asn Ser Glu |

| | | | | | |
|-----------------|-----|---------------------|-----|---------------------|-----|
| Leu Phe Glu Phe | 350 | Asp Thr Pro Gly | 355 | Leu Pro Pro Ser | 360 |
| 365 | Ala | 370 | Asn | 375 | Met |
| Asn Leu Ser Gln | 380 | Leu Leu Gly Leu Arg | 385 | Asn Ile Ser Ile | 390 |
| 395 | Leu | 400 | Ala | 405 | Val |
| Gln Ala Tyr Gln | 410 | Gln Cys Lys Glu Gly | 415 | Ala Leu Trp Thr | 420 |
| 425 | Leu | 430 | Glu | 435 | Ile |
| Leu Gln Leu Asn | 440 | Asp Ser Tyr Asp Leu | 445 | Glu Glu His Leu Asp | 450 |
| 455 | Asn | 460 | Leu | 465 | Lys |
| Asn Gln Tyr Thr | 470 | Asn Lys Leu Arg Gln | 475 | Leu Gln Ser Leu | 480 |
| 485 | Val | 490 | Ser | 495 | Arg |
| Val Asp Thr Gln | 500 | Ser Leu Asp Leu Leu | 505 | Ser Ala Ala Arg | 510 |
| 515 | Asp | 520 | Leu | 525 | Arg |
| Asp Leu Glu Ala | 530 | Leu Gln Ser Ser Gly | 535 | Gln Arg Ile His | 540 |
| 545 | Pro | 550 | Val | 555 | Tyr |
| Pro Asp Phe Leu | 560 | Val Gln Ile Gln Arg | 565 | Val Val Lys Thr | 570 |
| 575 | Met | 580 | Gly | 585 | Gln |
| Met Glu Gln Leu | 590 | Ala Gln Glu Leu Gln | 595 | Leu Ala Gln Ala | 600 |
| 605 | Asp | 610 | Gln | 615 | Gly |
| Asp Asn Ser Val | 620 | Leu Gly Gln Arg Leu | 625 | Glu Glu Ala Gln | 630 |
| 635 | Leu | 640 | Val | 645 | Leu |
| Leu Arg Asn Leu | 650 | His Gln Glu Lys Val | 655 | Val Pro Gln Gln Ser | 660 |
| 665 | Val | 670 | Ala | 675 | Ala |
| Val Ala Lys Leu | | Asn Leu Ser Val Arg | | Leu Glu Ser Ser | |
| 530 | Pro | 535 | Val | 540 | Thr |
| Pro Asn Leu Gln | 545 | Leu Glu Thr Ser Asp | 550 | Leu Ala Asn Val | 555 |
| 560 | Tyr | 565 | Ala | 570 | Arg |
| Tyr Leu Lys Gly | 575 | Glu Leu Pro Ala Trp | 580 | Ala Arg Ile Leu | 585 |
| 590 | Asn | 595 | Glu | 600 | Ser |
| Asn Val Ser Glu | 605 | Cys Phe Leu Ala Arg | 610 | Met Gly Tyr Phe | 615 |
| 620 | Gln | 625 | Val | 630 | Ala |
| Gln Tyr Val Ala | 635 | Trp Val Arg Glu Glu | 640 | Thr Gln Arg Ile | 645 |
| 650 | Thr | 655 | Asn | 660 | Ile |
| Thr Cys Gln Pro | 665 | Leu Ser Gly Ala Leu | 670 | Asn Ser Arg Val | 675 |
| 680 | Leu | | | | |
| Leu Cys Asp Met | | Met Ala Asp Pro Trp | | Asn Ala Phe Trp Phe | |
| 620 | Leu | 625 | Pro | 630 | Cys |
| Leu Ala Trp Cys | 635 | Thr Phe Phe Leu Ile | 640 | Ser Ile Ile Phe | 645 |
| 650 | Val | 655 | Ile | 660 | Ser |
| Val Lys Thr Ser | 665 | Lys Tyr Phe Arg Pro | 670 | Ile Arg Lys Arg Leu | 675 |
| 680 | Ser | | | | |
| Ser Thr Ser Ser | | Glu Glu Thr Gln Leu | | Phe His Ile Pro Arg | |
| 680 | Thr | | | | |
| Thr Ser Leu Lys | | Leu | | | |

<210> 33

<211> 590

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2793768CD1

<400> 33

| | | | | | | | | | | | |
|-----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met Ser Ser Ala | Cys | Asp | Ala | Gly | Asp | His | Tyr | Pro | Leu | His | Leu |
| 1 | 5 | 10 | | | | 15 | | | | | |
| Leu Val Trp Lys | Asn | Asp | Tyr | Arg | Gln | Leu | Glu | Lys | Glu | Leu | Gln |
| 20 | 25 | | | | | 30 | | | | | |
| Gly Gln Asn Val | Glu | Ala | Val | Asp | Pro | Arg | Gly | Arg | Thr | Leu | Leu |
| 35 | 40 | | | | | 45 | | | | | |
| His Leu Ala Val | Ser | Leu | Gly | His | Leu | Glu | Ser | Ala | Arg | Val | Leu |
| 50 | 55 | | | | | 60 | | | | | |
| Leu Arg His Lys | Ala | Asp | Val | Thr | Lys | Glu | Asn | Arg | Gln | Gly | Trp |
| 65 | 70 | | | | | 75 | | | | | |
| Thr Val Leu His | Glu | Ala | Val | Ser | Thr | Gly | Asp | Pro | Glu | Met | Val |

| | | | | | |
|-----------------|---------------------|---------------------|-----|--|----|
| | 80 | | 85 | | 90 |
| Tyr Thr Val Leu | Gln His Arg Asp Tyr | His Asn Thr Ser Met | Ala | | |
| | 95 | 100 | 105 | | |
| Leu Glu Gly Val | Pro Glu Leu Leu Gln | Lys Ile Leu Glu Ala | Pro | | |
| | 110 | 115 | 120 | | |
| Asp Phe Tyr Val | Gln Met Lys Trp Glu | Phe Thr Ser Trp Val | Pro | | |
| | 125 | 130 | 135 | | |
| Leu Val Ser Arg | Ile Cys Pro Asn Asp | Val Cys Arg Ile Trp | Lys | | |
| | 140 | 145 | 150 | | |
| Ser Gly Ala Lys | Leu Arg Val Asp Ile | Thr Leu Leu Gly Phe | Glu | | |
| | 155 | 160 | 165 | | |
| Asn Met Ser Trp | Ile Arg Gly Arg Arg | Ser Phe Ile Phe Lys | Gly | | |
| | 170 | 175 | 180 | | |
| Glu Asp Asn Trp | Ala Glu Leu Met Glu | Val Asn His Asp Asp | Lys | | |
| | 185 | 190 | 195 | | |
| Val Val Thr Thr | Glu Arg Phe Asp Leu | Ser Gln Glu Met Glu | Arg | | |
| | 200 | 205 | 210 | | |
| Leu Thr Leu Asp | Leu Met Lys Pro Lys | Ser Arg Glu Val Glu | Arg | | |
| | 215 | 220 | 225 | | |
| Arg Leu Thr Ser | Pro Val Ile Asn Thr | Ser Leu Asp Thr Lys | Asn | | |
| | 230 | 235 | 240 | | |
| Ile Ala Phe Glu | Arg Thr Lys Ser Gly | Phe Trp Gly Trp Arg | Thr | | |
| | 245 | 250 | 255 | | |
| Asp Lys Ala Glu | Val Asn Gly Tyr | Glu Ala Lys Val Tyr | Thr | | |
| | 260 | 265 | 270 | | |
| Val Asn Asn Val | Asn Val Ile Thr Lys | Ile Arg Thr Glu His | Leu | | |
| | 275 | 280 | 285 | | |
| Thr Glu Glu Glu | Lys Arg Tyr Lys | Ala Asp Arg Asn Pro | Leu | | |
| | 290 | 295 | 300 | | |
| Glu Ser Leu Leu | Gly Thr Val Glu His | Gln Phe Gly Ala Gln | Gly | | |
| | 305 | 310 | 315 | | |
| Asp Leu Thr Thr | Glu Cys Ala Thr Ala | Asn Asn Pro Thr Ala | Ile | | |
| | 320 | 325 | 330 | | |
| Thr Pro Asp Glu | Tyr Phe Asn Glu Glu | Phe Asp Leu Lys Asp | Arg | | |
| | 335 | 340 | 345 | | |
| Asp Ile Gly Arg | Pro Lys Glu Leu Thr | Ile Arg Thr Gln Lys | Phe | | |
| | 350 | 355 | 360 | | |
| Lys Ala Met Leu | Trp Met Cys Glu Glu | Phe Pro Leu Ser Leu | Val | | |
| | 365 | 370 | 375 | | |
| Glu Gln Val Ile | Pro Ile Ile Asp Leu | Met Ala Arg Thr Ser | Ala | | |
| | 380 | 385 | 390 | | |
| His Phe Ala Arg | Leu Arg Asp Phe Ile | Lys Leu Glu Phe Pro | Pro | | |
| | 395 | 400 | 405 | | |
| Gly Phe Pro Val | Lys Ile Glu Ile Pro | Leu Phe His Val Leu | Asn | | |
| | 410 | 415 | 420 | | |
| Ala Arg Ile Thr | Phe Gly Asn Val Asn | Gly Cys Ser Thr Ala | Glu | | |
| | 425 | 430 | 435 | | |
| Glu Ser Val Ser | Gln Asn Val Glu Gly | Thr Gln Ala Asp Ser | Ala | | |
| | 440 | 445 | 450 | | |
| Ser His Ile Thr | Asn Phe Glu Val Asp | Gln Ser Val Phe Glu | Ile | | |
| | 455 | 460 | 465 | | |
| Pro Glu Ser Tyr | Tyr Val Gln Asp Asn | Gly Arg Asn Val His | Leu | | |
| | 470 | 475 | 480 | | |
| Gln Asp Glu Asp | Tyr Glu Ile Met Gln | Phe Ala Ile Gln Gln | Ser | | |
| | 485 | 490 | 495 | | |
| Leu Leu Glu Ser | Ser Arg Ser Gln Glu | Leu Ser Gly Pro Ala | Ser | | |
| | 500 | 505 | 510 | | |
| Asn Gly Gly Ile | Ser Gln Thr Asn Thr | Tyr Asp Ala Gln Tyr | Glu | | |
| | 515 | 520 | 525 | | |
| Arg Ala Ile Gln | Glu Ser Leu Leu Thr | Ser Thr Glu Gly Leu | Cys | | |
| | 530 | 535 | 540 | | |
| Pro Ser Ala Leu | Ser Glu Thr Ser Arg | Phe Asp Asn Asp Leu | Gln | | |
| | 545 | 550 | 555 | | |
| Leu Ala Met Glu | Leu Ser Ala Lys Glu | Leu Glu Glu Trp Glu | Leu | | |
| | 560 | 565 | 570 | | |
| Arg Leu Gln Glu | Glu Glu Ala Glu Leu | Gln Gln Val Leu Gln | Leu | | |
| | 575 | 580 | 585 | | |

Ser Leu Thr Asp Lys
590

<210> 34
<211> 315
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3035248CD1

<400> 34
Met Val Gly Pro Trp Val Tyr Leu Val Ala Ala Val Leu Leu Ile
1 5 10 15
Gly Leu Ile Leu Phe Leu Thr Arg Ser Arg Gly Arg Ala Ala Ala
20 25 30
Ala Asp Gly Glu Pro Leu His Asn Glu Glu Arg Ala Gly Ala
35 40 45
Gly Gln Val Gly Arg Ser Leu Pro Gln Glu Ser Glu Glu Gln Arg
50 55 60
Thr Gly Ser Arg Pro Arg Arg Arg Arg Asp Leu Gly Ser Arg Leu
65 70 75
Gln Ala Gln Arg Arg Ala Gln Arg Val Ala Trp Glu Asp Gly Asp
80 85 90
Glu Asn Val Gly Gln Thr Val Ile Pro Ala Gln Glu Glu Glu Gly
95 100 105
Ile Glu Lys Pro Ala Glu Val His Pro Thr Gly Lys Ile Gly Ala
110 115 120
Lys Lys Leu Arg Lys Leu Glu Glu Lys Gln Ala Arg Lys Ala Gln
125 130 135
Arg Glu Ala Glu Glu Ala Glu Arg Glu Glu Arg Lys Arg Leu Glu
140 145 150
Ser Gln Arg Glu Ala Glu Trp Lys Lys Glu Glu Glu Arg Leu Arg
155 160 165
Leu Lys Glu Glu Gln Lys Glu Glu Glu Glu Arg Lys Ala Gln Glu
170 175 180
Glu Gln Ala Arg Arg Asp His Glu Glu Tyr Leu Lys Leu Lys Glu
185 190 195
Ala Phe Val Val Glu Glu Glu Gly Val Ser Glu Thr Met Thr Glu
200 205 210
Glu Gln Ser His Ser Phe Leu Thr Glu Phe Ile Asn Tyr Ile Lys
215 220 225
Lys Ser Lys Val Val Leu Leu Glu Asp Leu Ala Phe Gln Met Gly
230 235 240
Leu Arg Thr Gln Asp Ala Ile Asn Arg Ile Gln Asp Leu Leu Thr
245 250 255
Glu Gly Thr Leu Thr Gly Val Ile Asp Asp Arg Gly Lys Phe Ile
260 265 270
Tyr Ile Thr Pro Glu Leu Ala Ala Val Ala Asn Phe Ile Arg
275 280 285
Gln Arg Gly Arg Val Ser Ile Thr Glu Leu Ala Gln Ala Ser Asn
290 295 300
Ser Leu Ile Ser Trp Gly Gln Asp Leu Pro Ala Gln Ala Ser Ala
305 310 315

<210> 35
<211> 2345
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1889577CB1

<400> 35

```

ggagccgggc ccgagcacca ggcgcaggcc cggcgcccg cgtcccgcac cctcgtcctc 60
acagacgcca cagccatggc catgatggtg tttccgcggg aggagaagct gagccaggat 120
gagatcgtgc tgggcaccaa ggctgtcatc cagggactgg agactctgcg tggggagcat 180
cgtgccctgc tggctcctct ggttgacact gaggcggcg aacccgagcc tggctcgcag 240
gagcgctgca tcctcctgcg tcgctccctg gaagccattg agcttgggct gggggaggcc 300
caggtgatct tggcatgtc cagccactg ggggctgtag aatcagagaa gcagaagctg 360
cgggcgcagg tgcggcgtct ggtgcaggag aaccagtggc tgcgtgagga gctggcgggg 420
acacagcaga agctgcagcg cagtgcagag gccgtggccc agctcgagga ggagaagcag 480
cacttgctgt tcatgagcca gatccgcaag ttggatgaag acgcctcccc taacgaggag 540
aaggggggac tccccaaaga cacactggat gacctgttcc ccaatgagga tgagcagag 600
ccagccctta gcccgaggag aggggatgtg tctggtcagc atgggggcta cgagatcccg 660
gcccggtccc gcacctgca caacctggtg atccaatag cctcacaggg ccgtacagag 720
gtagctgtgc cactctgcaa gcaggcactc gaagacctgg agaagacgtc aggccacgac 780
cacctgagc ttgccaccat gctgaacatc gctcactgg tctatcgga tcagaacaag 840
tacaaggagg ctgcccacct gctcaatgat gctctggcca tccgggagaa aacactgggc 900
aaggaccacc cagcgtggc tgcgacacta aacaacctgg cagtcctgta tggcaagagg 960
ggcaagtaca aggaggtga gccattgtgc aagcgggcac tggagatccg ggagaaggtc 1020
ctgggcaagt ttaccacaga tgtggccaag cagctcagca acctggccct gctgtgcag 1080
aaccagggca aagctgagga ggtggaatat tactatcggc gggcactgga gatctatgct 1140
acacgcctcg ggccgatga tccaatgtg gccaaagcca agaacaacct ggcttctcgc 1200
tacctgaagc agggcaagta ccaggatgcg gagaccttgt acaaggagat agcccatctg gatgcacgca 1320
gctcatgaga aagagtgtg ggataagcgc cgggacagcg cccctatgg ggaatacggc 1380
agctgggtaca aggcctgtaa agtagacagc cccacagtca acaccacct gcgcacgttg 1440
ggggccctat accggcgcca gggcaagctg gaagccgcgc acacactaga ggactgtgcc 1500
agcgttaacc gcaagcaggg tttggacccc gcaagccaga ccaaggtggt agaactgctg 1560
aaagatggca gtggcaggg gggagacccc cgagcagcc gagacatggc tgggggtgcc 1620
gggcctcggt ggtgctctga cctcaggac gtgggacct cagctgagtg gaatggggat 1680
ggcagtggt ccttgaggcg cagcggttcc tttgggaaac tccgggatgc cctgaggcgc 1740
agcagtgaga tgctgtaaaa gaagctgcag gggggcacc ccaggaggcc ccctaacccc 1800
aggatgaagc gggccagttc cctcaacttc ctcaacaaga gcgtggaaga gccgaccag 1860
cctggaggca caggtctctc tgacagccgc actctcagct ccagctccat ggacctctcc 1920
cgacgaagct cctggtggg ctaatgctga aggggcagcc agtcaccaga gcgccacct 1980
ggcacacccc cctcacccca gccctgcgca tgggocctgt gcttgteccg cctgtctctc 2040
ccacagcccc tgtctttct gttcaatctc agggtaacct tctcccttgt catctcagcc 2100
tgagccctgg agcgtggcc tgcccactcc agctccatcc cttatttatt ccttcagca 2160
gggcctctt ccctaggttc gggccagcag gaggtgccg ctggagtctc caccatagac 2220
tcagtggcct ggcctcccca gacccagag ccaagaacac taagcactcg ccggcccttc 2280
ggcacccctg cctcctctc cgaactcaac cgccggttgc ttctgtatat agagaaataa 2340
gttat 2345

```

<210> 36

<211> 709

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2427982CB1

<400> 36

```

gtgacagtag caaccgccg aatggcgaaa gcaacaacaa tcaaagaagc cttagcgaga 60
tggaagaga aaactggcca gaggccatct gaagccaaag agataaaaact ttatgccag 120
attcccccta tagagaagat ggtatgcattc ttgtccatgc ttgctaattg cgagaagctt 180
tactgtctca caaactgcat tgaaaaaatt gccaacctga atggcttaaa aaacttgagg 240
atattatctt taggaagaaa caacataaag aacttaaatg gactggaggc agtaggggac 300
acattagaag aactgtggat ctctacaat tttattgaga agttgaaagg gatccacata 360
atgaagaaat tgaagattct ctacatgtct aataacctgg taaaagactg ggctgagttt 420
gtgaagctgg cagaactgcc atgctcgaa gacctggtgt ttgtaggcaa tcccttgga 480
gagaacatt ctgctgagaa taactggatt gaagaagcaa ccaagagagt gcccaaactg 540
aaaaagctgg atggtactcc agtaattaaa ggggatgagg aagaagacaa ctaatgccac 600
gctttccact gtgtgttaac ttatttaaat gtcataagaa attttatata 660
attgtctatt ttaaagattc tgtatgggac aaaagtttct taagataaa 709

```

<210> 37

<211> 1569

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2470833CB1

<400> 37

| | | | | | | |
|------------|------------|-------------|-------------|-------------|-------------|------|
| ctcgaccgct | cgagccgaat | tcggctcgag | attgagttta | agtggcctgc | agaaccccg | 60 |
| gggcccaggc | ggcacggcgg | gcgctggggc | tcggggcaga | gcttttcagga | ggtttatggc | 120 |
| agcttcactt | tcacggcatc | cacccccctc | gggcccctgc | gcagagagga | ggaagctcct | 180 |
| gcggctgag | cgggectgga | ggaagtgagc | agcggggctc | ctgcctccc | gcctggtccc | 240 |
| cgaagacccc | agaagaaccc | ggaacttgct | tccattcgga | atccagggac | caccctttgc | 300 |
| actcagtagg | cctttgtttt | cctgcgtgga | aagcggttgg | gcttggggagg | cgatggagcc | 360 |
| ggagttcttg | tacgacctgc | tgcagctccc | caaggggggtg | gagccccag | cggaggagga | 420 |
| gctctcaaaa | ggaggaaa | agaaatacct | gccacccact | tcccgggaagg | accccaaatt | 480 |
| tgaagaactg | cagaaggtgt | tgatggagtg | gatcaatgcc | actcttctcc | ccgagcacat | 540 |
| tgtggtccgc | agcctggagg | aggacatgtt | cgacgggctc | atcctacacc | acctattcca | 600 |
| gaggctggcg | gcgctcaagc | tgggaagcaga | ggacatcgcc | ctgacagcca | caagccagaa | 660 |
| gcacaagctc | acagtgggtg | tggaggccgt | gaaccggagt | ctgcagctgg | aggagtggca | 720 |
| ggccaagtgg | agcgtggaga | gcattcttcaa | caaggacctg | ttgtctaccc | tgcacctcct | 780 |
| tgtggccctg | gccaagcgct | tccagcccga | cctctccctc | ccaaccaacg | tccaggtgga | 840 |
| ggtcatcact | atcgagagca | ccaaaagtgg | tctgaagtca | gagaagttgg | tggaaacagct | 900 |
| cactgaatac | agcacagaca | aggacgagcc | tccaaaggac | gtctttgatg | aattatttaa | 960 |
| gctggctccg | gagaaagtga | acgcagtga | agaggccatc | gtgaactttg | tcaaccagaa | 1020 |
| gctggacgcg | ctgggectgt | ctgtgcagaa | tctggacacc | cagtttgtag | atggggtcat | 1080 |
| cttactcttg | ctgattggac | aacttgaagg | cttcttcctg | cacttaaagg | aattctacct | 1140 |
| cactcccaac | tctcctgcag | aaatgctgca | caacgtcacc | ctggcgctgg | agctgctgaa | 1200 |
| ggacgagggc | ctgctcagct | gccctgtcag | ccttgaagat | atcgtgaaca | aggatgcca | 1260 |
| gagcacactg | cgggtgctct | atgggtctgt | ctgcaagcac | acgcagaagg | cacacaggga | 1320 |
| caggacgccc | catggagccc | cgaattgacc | ctcactgcct | ccaaagccca | gagcctgcct | 1380 |
| gtcagcccag | ctggaggggc | cgaggctgca | gggtgtcctc | ccacagtccc | gctgtttcct | 1440 |
| gtgcattcgt | gaccgccttc | cctcccacc | tgtctcctgt | ctccatcgtt | ggattatctt | 1500 |
| tgaacccctt | tgtgtggatc | attttgagcc | gcctggcctt | gctcagttta | ttttaataaa | 1560 |
| agtatttct | | | | | | 1569 |

<210> 38

<211> 1172

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2080579CB1

<400> 38

| | | | | | | |
|------------|------------|-------------|-------------|------------|-------------|------|
| cgacgggctg | cggcctgcgg | aacctgaggc | agctggggag | ggccggggcg | gccggccgga | 60 |
| tagcgagccg | cgctggcggc | ggcgggtggc | gcgatgatgg | agatccagat | ggacgagggc | 120 |
| ggcggcgtgg | tggtgtacca | ggacgactac | tgctccggct | cggtgatgtc | ggagcgggtg | 180 |
| tcgggcctgg | cgggctccat | ctaccgcgag | ttcgagcgcc | tcatccactg | ctacgacgag | 240 |
| gaggtgggtc | aggagctcat | gccgctgggt | gtgaacgtgc | tggagaacct | agactcgggt | 300 |
| ctcagcgaga | accaggagca | cgaggtggag | ctggagctgc | tgcgcgagga | caacgagcag | 360 |
| ctgctcacc | agtacgagcg | tgagaaggcg | ctgcgcaggg | aggcggagga | gaaattcatt | 420 |
| gagtttgaag | atgctctgga | acaagagaag | aaagagctgc | aatccagggt | ggagcactac | 480 |
| gagttccaga | cgcgcagct | ggagctgaag | gccaagaact | atgccgatca | gatttccggg | 540 |
| ttggaggagc | gggagtcgga | gatgaagaag | gagtacaatg | ccctgcacca | gcggcacaca | 600 |
| gagatgatac | agacctacgt | ggagcacatt | gagagggtcca | agatgcagca | ggtcggagga | 660 |
| aacagccaga | cagagagcag | cctgccgggg | cggagggtacg | cggggcgcg | cgggggtggag | 720 |
| gtacgcgggg | cgcggcgggg | cggagggtacg | caggacgcgg | cacatgccag | ggtcgtagtg | 780 |
| cttgttatgg | ccgcgctct | gggctcagga | tgacagggaa | cactctggag | acccaggagg | 840 |
| agggagctgg | ttggagcgtg | gctgagcagg | gatgtggggg | ggcggccctg | ggcgacggac | 900 |
| atgtgtcgag | cccgtgtgtc | ccttctgtct | ggtccacaat | tcagtcaagg | aggccccggg | 960 |
| cagcttgctc | atatttaacc | cttactatcc | acagattctg | tgtttgcaaa | tgtatctact | 1020 |
| cgctaaaata | tatctgtgac | ctcagcggg | cacagtgaac | tatgcctgta | atcctagaac | 1080 |
| tttgggaggg | caaggcagga | ggattacttg | agcccaggag | ttcgagctca | gcttgggcaa | 1140 |
| ccagcaagat | cttgtctcta | caaaaaaaaa | aa | | | 1172 |

<210> 39

<211> 2380
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2156553CB1

<400> 39
cagggtttca ctgtcttagc caggatgggc tcgatctcct gacctcatga tccatccacc 60
tcagcctccc aaagtgtctgg gattacaggc gtcagccacc gtgaccagcc tgaacaggga 120
gcaagttcta aactcaggct ctagagtcag aaaaggtaga gtcaggttct ggatccaaaa 180
tggggcaagt catgatcagg ttctggaacc agaacaggcc tcaagcctag gggctgagca 240
gggtatcccc tggcctggga gcagaggact tctggctgac tgctgcccg aacgtttctca 300
agctgggtgg gaagtctgag tgggaagtcac accctattca ggcagtagag gaagaggcct 360
caggagacaa gcagcccaag aaacaggaga aaaaccaggt gttgggtgtcc ccagagtttg 420
tggatgaagc tctgtgtgcg tgcgaggagt accttagcaa cttggccacc atggacatcg 480
acaaggacct ggaggccccc ctgtacctca ccccgagggy ctgggtccctc ttcctccagc 540
gctactacca agtgggtccac gaagggggcag aactcaggca cctcgacact cagggtccagc 600
gctgtgagga catcctgcag cagctgcagg ccgtgggtacc ccagatagac atggaagggg 660
atcgcaacat ctggatcgtg aagccaggag ccaagtcctg tggacgaggg atcatgtgca 720
tggaacacct ggaggagatg ctgaagctgg tgaacggcaa ccccggtggg atgaaggacg 780
gcaagtgggt ggtgcagaag tatattgagc ggccccctct catctttggc accaagtttg 840
acctcagaca cctgttctct gtaactgact ggaaccact taccgtgtgg tctaccggc 900
acagctatat cgcgttttcc acgcagccct tctccctgaa gaacctggac aactcagtgc 960
acctgtgcaa caactccatc cagaagcacc tggagaactc atgccatcgg catccactgc 1020
ttccgccaga caacatgtgg tctagccaga ggttccaggc ccacctgcag gagatgggtg 1080
ccccaaatgc ttggtccacc atcatcgtgc ctggcatgaa ggatgctgtg atccacgcac 1140
ttcagacctc ccaggacacc gtgcagtgtc ggaaggccag ctttgagctc tatggcgctg 1200
acttcgtgtt cggggaggac ttccagccct ggctgattga gatcaacgcc agccccacga 1260
tggaacacct cacagcagtc actgcccggc tctgtgtctg cgtgcaagct gacaccctgc 1320
gcgtgggtcat tgaccggatg ctggaccgca actgtgacac aggagccttt gagctcatct 1380
ataagcagcc tgcgtgtgag gtgcctcaat atgtgggcat cgggtcctg gtagagggct 1440
tcaccatcaa gaagcccatt gcgatgtgtc atcgccggat gggggtccgc ccagcagtc 1500
ctctgtctgac ccagcagggc tctggggaag gcaaggactc ggggatccct acccacaggt 1560
cagcttcttag gaaaggcact gggggccagga gcctggggca cagtgagaag ccagtctcca 1620
ctgcaccac ttcagccccc ggaaagggga agaaaggcaa ggcgaaaagg gccacagccc 1680
tggtctgccc caatctctgg gagtgggatg cccccagcac caggatgggc tgcattttca 1740
ccatgacctt ttctagtggg gacaggcaac cccaccactt gaacagattg ccactgagtc 1800
cgaagaacct ccaggccctg ggtaagacca ttcccccaaa acacccgagt gttccaagge 1860
gatttatccc tgcttccag gccctcccca accacctgga tcagccacc caccaaagg 1920
ccaccagtag caagtaaaag ccaactactca caaagtattg tttaaaaaa cacagccaaa 1980
ttagctgggc acggtggtgt gagcctgtgg tccaggctac tgcgagggt aatgaggatc 2040
gcttgagccc gggaggtcaa ggctgcagtg agttatgatc gcacctctgc actccagcct 2100
gggagacaga gcgagatgct gtcttaaaaa caaaaacaca aaaaaagcac tttgggaggc 2160
cgaggcgggc agatcacttg aggtcaggag ttcgagacca gcctggccaa catggtgaaa 2220
ccccatctct actaaaaaat acaaaaatta gccgagtgtg ggggcagggt cctataatcc 2280
cagctacttg ggagactgag gcacgagaat cgcttgaacc caggaagcgg aggttgagc 2340
gagccgagat tgcaccattg cactccagca tgggcagcat 2380

<210> 40
<211> 4396
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2182855CB1

<400> 40
tgccgctgca gggggctaga gatgcctcgt ggggttttcc agcagctttc caacttggtt 60
ttgcagggaac ttaacgcca cctgagcaac ctaacctcag cgtttgaaaa agcaacagct 120
gagaaaaatca agtgtcagca agaggccgat gccacgaaca ggtgatctt actggcgaa 180
aggctggtcg ggggattagc atccgctggg ctgagctctg ggagaacttc 240
aggagccagg gggtcacgct gtgtggggac gtccctgctc tctctgcctt cgtgtcctac 300
gtgggctact tcaccaagaa ataccggaat gagctgatgg agaaattctg gateccttac 360
atacataact taaaggtccc catcccgatc acgaatggcc tggatccctt gagcctgctg 420

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| acagatgacg | cggacgtggc | cacctggaac | aaccaggggc | tccccagcga | ccgcatgtcc | 480 |
| accgagaatg | ccaccatcct | gggcaacacc | gagcgggtgg | cgctgatacgt | ggagcggccag | 540 |
| ctccaaggaa | tcaagtggat | caaaaacaaa | tacaggagtg | aactgaaagc | catccgcctg | 600 |
| ggacagaaga | gctacctgga | tgtcatcgag | caggccatct | cggaagggga | cacgttgctc | 660 |
| attgagaaca | tggcgaaac | cgtaggaccc | gtgctggacc | ctctactggg | caggaacacg | 720 |
| attaaaaagg | gaaagtacat | taagatcggg | gacaaggagg | tggagtacca | ccccaggttc | 780 |
| cgctgatcc | tacacaccaa | gtacttcaac | ccacactaca | agccagagat | gcaggctcag | 840 |
| tgcacctca | tcaacttcct | ggtcaccagg | gatggactcg | aggaccaact | cttggcgcgt | 900 |
| gtggtggcca | aagagcggcc | agatctggaa | cagctgaagg | caaacctcac | caagtcccaa | 960 |
| aacgaattta | agattgttct | gaaagagctg | gaagactcgc | tcctggcccg | tctgtcggct | 1020 |
| gcgtcgggga | actttctggg | agacacggcc | ttggtggaga | atctggagac | caccaagcac | 1080 |
| acagccagcg | agatcgagga | gaaggtgggtg | gaggcaaaaa | tcacagaagt | taaaatcaac | 1140 |
| gaagcgagag | agaactaccg | cccggtctcg | gagagggcac | ctctgtctta | cttcatactg | 1200 |
| aacgatctca | acaaaaatcaa | ccccgtctac | cagttctccc | tcaaggcctt | caacgtgggtg | 1260 |
| tttgacggga | acttccagag | gaccaccctc | gccaacgagg | tgaagcagcg | ggtgtgggag | 1320 |
| ctgacggagc | agatcaccta | ctccgtctac | atgtacacgg | ccgggggact | cttcgagagg | 1380 |
| gacaaactca | ttttctggc | acaagttacg | tttcaggtcc | tgtccatgaa | gaaggagctg | 1440 |
| aaccagtg | agctggattt | cctcctgcgg | ttcccttcta | aggccggagt | ggtctcacca | 1500 |
| gtggacttcc | tccagcatca | aggctggggc | gggatacaagg | ccctctcgga | gatggatgag | 1560 |
| ttcaaaaaat | tggacagtga | catcgaaagg | cttgccaagc | gctggaaaaa | gctgggtggag | 1620 |
| tcggaagccc | ccgagaagga | gatcttcccc | aaggagtggg | agaacaagac | ggccctgcag | 1680 |
| aagctgtgca | tgggtcgctg | cctgcggcca | gatcgcatga | cctacgctat | caagaacttc | 1740 |
| gtggaggaaa | agatgggcag | caagttcgtg | gaaggccgga | gtgttgagtt | ttctaagtc | 1800 |
| tacgaggaga | gcagccctc | cacgtcaatc | ttcttcatcc | tctccccggg | ggttgacccc | 1860 |
| ttgaaagacg | tggaaagccct | gggaaaaaaa | ctagggttta | ccatagacaa | tggaaaactc | 1920 |
| cataatgtgt | ccctggggca | gggacaagag | gtggtggctg | agaacgccct | ggacgtgggt | 1980 |
| gcagagaaa | gacactgggt | cattctgcag | aatatccacc | tgggtggccc | gtggctggga | 2040 |
| acactggaca | agaagctgga | gcgtacagc | acgggcagcc | atgaggacta | ccgggtgttc | 2100 |
| atcagcgagg | agcctgcccc | cagtcctcag | acccacatca | tccccagg | cattctggag | 2160 |
| aacgccatca | agatcaccaa | cgagccccc | acgggcatgt | acgccaaact | gcacaaggcc | 2220 |
| ctggacctgt | tcaccagga | caccctggag | atgtgcacca | aggagatgga | gttcaagtgc | 2280 |
| atgctcttgc | ccctgtgcta | cttccacgct | gtggtggcag | agaggcgcaa | gttcggcgcc | 2340 |
| cagggtctgga | accgtctgta | cccttcaac | aacggggacc | tcaccatctc | catcaacgtg | 2400 |
| ctctacaact | acctggaggc | caaccccaag | gtgccctggg | acgatctccg | ctaccttttt | 2460 |
| ggtgaaatca | tgtatggcgg | ccacatcaca | gatgactggg | accgtcggtt | gtgcaggacc | 2520 |
| tacctggctg | aatacatccg | gacggagatg | ctggaggagg | acgtcctgct | ggccccgggc | 2580 |
| tttcagatcc | cccccaacct | ggactacaag | ggttaccacg | aatacatcga | tgagaacctg | 2640 |
| ccccctgaga | gtccctatct | gtatggcctg | caccccaacg | cagagatttg | ctttctgacg | 2700 |
| gtcacctcag | agaagctgtt | ccgcactgtc | ctggaaatgc | agccaaaaga | gacggactcg | 2760 |
| ggggcaggga | ggggagtgtc | ccgcgaggag | aagggtgaagg | ccgtgctgga | cgacatcctg | 2820 |
| gagaagatcc | cggagacttt | caacatggct | gagatcatgg | caaaggcagc | ggaaaagacc | 2880 |
| cctacgtgg | tagtgcctt | tcaagaatgt | gaaagaatga | acatcctgac | caacgaaatg | 2940 |
| cgccgttcgc | tcaaggagct | gaacctgggg | ctgaaggagg | aactgaccat | cacgaccgac | 3000 |
| gtggaagatc | tgtccacggc | tctcttctat | gacaccgtgc | ctgatacgtg | ggtggcccgg | 3060 |
| gcctacccct | ccatgatggg | cctggcggcc | tggtacgcag | acctgtctgt | ccgcatcagg | 3120 |
| gaactcaggg | catggacgac | agactttgct | gtcccacca | ccgtgtggct | ggccggcttc | 3180 |
| ttcaaccccc | agtcgttctt | cacggccatc | atgcagtcca | tggccaggaa | gaacgagtgg | 3240 |
| ccctgggaca | agatgtgtct | gtctgtcgag | gtgaccaaga | aaaaccgaga | ggacatgacc | 3300 |
| gtccccccac | gagagggtct | ctacgtgtac | ggactcttca | tggaaagggc | tcgctgggac | 3360 |
| accagactg | gagtcacgc | tgaagcgcg | ctgaaagagc | tgaccccgcc | catgcctgtc | 3420 |
| atcttcatca | aggccattcc | tgtggaccgc | atggagacca | agaacatcta | tgagtgtccc | 3480 |
| gtgtacaaaa | cacgcacccg | cgccccacc | tatgtctgga | cctttaactt | gaagacaaa | 3540 |
| gagaaggcag | cgaagtggat | cctggcagcc | gtggcgctgc | tcctacaggt | ttagctcgct | 3600 |
| cctgcctcac | agcccacact | ccctggggct | ggaccacaac | tcagcccttc | acctgtgcac | 3660 |
| ctgtgactta | ttctttacag | gaactgggtg | tggttttttc | ttctcttaaa | taatcagggtg | 3720 |
| ctttgtaacc | aagcacatcg | gaaccagagg | gtggagggtg | gtgtggaaga | ggtggggcag | 3780 |
| attaaagcca | gtggagccac | tcagctgtgc | ccatccattc | tgtgcctgat | ggccactgtg | 3840 |
| aggcctgggt | caggcttttg | ggaaagggcc | caattccag | cagccagagg | caagcatttc | 3900 |
| aggaagttaa | ttccaaatcc | tgacttcccg | ggggggtcag | tggagtattgc | ttttaacagg | 3960 |
| agcagcatgt | ggggtgaggg | cttggcactc | acaccatctg | gcctgatgtc | tcatgcaaac | 4020 |
| cacctgccaa | tgaggggaca | gggcgttacc | ccataccagg | agctgcccgc | tccacacagc | 4080 |
| gtctcatccg | ccattactcc | ctgctcccaa | aaacaacagg | acctttttaca | agtaaggaaa | 4140 |
| cacacccaag | ccccggcccc | cagagtctag | aagaacaagc | cgttcccaca | agcaagcccc | 4200 |
| ggacctgtag | gactctgtgg | ctgggctgat | gcaggacagc | gcagtgagggt | ggccaggcta | 4260 |
| gctagggcag | cagggtgcag | gcacgagggg | agagaattaa | ctagcaacct | gggggatgag | 4320 |
| gggagacagc | ggcaaacgct | agaaaagacg | gatcccaggg | ggaccaagac | aggatcaaca | 4380 |
| tttttttttt | tttttcg | | | | | 4396 |

<210> 41
 <211> 1831
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2242106CB1

<400> 41
 gtgagctgcc tgctgtgtcc gcctcctctc ctctcaggac tcagcttccc gccctgcaga 60
 gggcacaagg gagctttcca tgggcccttt ggcccggaga caggaaggcg ggccctgacc 120
 ctctcagtct ggagcccagg tgaaggtctg cagttagccc tgaggggtgg aactcggggg 180
 aggagtctgg aagggggcca gggtcggggc caggactggg ggctgtggag aggccagaag 240
 aggcaggtat ccttgggaag ccgtcagtgg tagggggcag gcaacgctgg agggagctgc 300
 cagggtcagg gcaggggtgt ccggcaaagc tccctggggg acaacagggg agctgggggc 360
 taggactctc cctctgtggg gatgacttcc agccaagtgg ctctgcccgc tggccctggc 420
 cctggccctg taggtgggct ggggagggtg tttggccagg aagtcaggct ggggtggggg 480
 ctgagcctct gccaggtcc actgtccctg caggagggtc gtacagtggg gaggtgttgg 540
 tgtggggacct gagccgtctt gaggaccgcg tgctgtggcg cacaggcctg acggatgaca 600
 cccacacaga ccctgtgtcc caggtcaggg caggggctgc cgggagggtt gggccctgag 660
 gcagagctgg ggagaggcgg cccactctcc taggaaggct acccgggcgt tggcagttgg 720
 acagctgcct cctgcattct cggagctccc gcagcatgag tcacccccct caggagccac 780
 tcccagggcc ctgtggccat tcccagcctc ctctgcaggc accgagttag cagcgtctggg 840
 attctggggac tggagagggc cctgggaacc ctccaggcct gccccctggg gggtagacct 900
 ggttctgggg cctcccggag aatttttttt tttcctggaa gagagggagg gttaggggtg 960
 agcgtgacac ctgggcaggg gtcccttgtc tccatcctgg ccttgcctgc tgttaactca 1020
 ggtggtgtgg ctggccgagc ctgggcacag ccaccgcttc cagggtgctga gtgtggccac 1080
 tgacggggaag gtgctactct ggcagggcct cggggtaggc cagctgcagc tcacagaggg 1140
 ctccgccctg gtcatgcagc agctgccacg gagcaccaag ctcaagaagc atccccggcg 1200
 ggagaccgag gtgggcgcca cggcagtggt cttctccagc tttgacccta ggctgttcat 1260
 tctgggcacg gaaggcggct tcccgcctca gtgttccctg gcagctggag aggcagccct 1320
 cagcgggatg cccagctccg tgcccctgcg gggcccagca cagtttacct tctccccca 1380
 cggcgggtccc atctactctg tgagctgttc ccccttccac aggaatctct tctgagcgc 1440
 tgggactgac gggcatgtcc acctgtactc catgctgcag gccctccctt tgacttcgct 1500
 gcagctctcc ctcaagtatc tgtttgtgtg gcgctgggtc ccagtgcggc ccttggtttt 1560
 tgagctgtcc tctgggaaag gtgacgtgca gctgtttgat ctccagaaaa gctcccagaa 1620
 acccacagtt ttgatcaagc aaaccagga tgaaagccct gtctactgtc tggagttaa 1680
 cagccagcag actcagctct tggctgcggg cgatgcccag ggcacagtga aggtgtggca 1740
 gctgagcaca gagtacacgg aacaagggcc ccgggaagct gaggacctgg actgcctggc 1800
 agcagaggtg gcggcctgag gggctccggg a 1831

<210> 42
 <211> 3249
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2726877CB1

<400> 42
 tttttatoga actcttatca ccttgttggc ccatagtagt ttaactgtgg ttgtgtttgc 60
 actttcaata ttatccagtt tgacattaaa tgaagagggt ggggaaaagc tattccatgc 120
 tcgaaacatt catcagactt ttcaactaat atttaataat ctcataaacg gtgatggcac 180
 tctaactaga aagtattcag ttgacctact gatggatctc cttaagaatc ctaaaattgc 240
 tgattatctc accagatatg agcacttttc ttcattgtct caccaagtat taggtcttct 300
 taatggaaag gatcctgatt cctcttcaaa ggttttagaa ttacttcttg ccttctgttc 360
 agtgactcag ctgcgccata tgctcactca gatgatgttt gaacagtctc cacctggcag 420
 cgccactctg ggaagccata ctaaatgttt agaacctact gtggctctac tgcgtgggtt 480
 aagccaacct ttggacggat cagaaaactg ttctgtttta gcattggagt tgttcaaggga 540
 aatatattgag gatgtcatag atgtctgctaa ctgttccctg gctgatcgtt ttgtgacctt 600
 tctgtgcctt acaatccttg atcaacttca gttcacagaa caaaatctag atgaggtttt 660
 aacaagacaa aaatgtgaaa ggattgcaa ggcctttgaa gttttgttaa ctctctgtgg 720
 agatgatata ctaaaaatgc atattgcaaa aatcttgaca actgtcaagt gtaccactct 780
 tatagaacaa caatttacat atggcaagat tgacctggga tttggaacaa aggttgacga 840
 ttctgaatta tgcaaaactg ctgctgatgt aattttgaaa actcttgatt tgattaacaa 900

```

acttaaacca ttggttctctg gtatggaagt aagcttctac aaaatacttc aggacccacg 960
tttgattact ccttttgctt ttgctttaac gtcagataat agagaacaag tacagtctgg 1020
actgagaata ttattggagg ctgctccact gccagatttt cctgctttag tacttggaaga 1080
aagtatagca gcaaaccaatg cctatagaca acaggaaaca gaacatatac ccagaaaaat 1140
gccctggcaa tcatcaaatc acagttttcc aacatcaata aagtgtttaa ctctcattt 1200
gaaagatggg gttcctggat tgaatattga agaatttaata gagaaacttc agtctggaat 1260
ggtggtaaac gatcagattt gtgatgtgag aatatctgac ataattggatg tatatgaaat 1320
gaaactatcc acattagctt ccaagaaag caggctacaa gatcttttgg aaacaaaagc 1380
tctagccctt gcacaggctg atagactgat tgctcagcat cgctgtcaaa gaactcaagc 1440
tgaacacagag gcacggacac ttgctagtat gttgagagaa gttgagagaa aaaatgaaga 1500
gcttagtggt ttgctgaagg cgcagcaagt tgaatcagaa agagcgcaga gtgatattga 1560
gcatctcttt caacataata ggaagttaga gtctgtggct gaagaacatg aaactactgac 1620
aaaaatcctac atggaacttc ttccagagaaa tgaaagtact gaaaagaaga ataaagattt 1680
acagatcaca tgtgattctc tgaataaaca aattgagaca gtgaaaagt tgaatgagtc 1740
actcaaggaa caaaatgaaa aaagtattgc ccaatttaata gagaaagaag aacagagaaa 1800
agaagtacag aatcagctag tagacagaga acataagcta gcaaatttgc atcaaaaaac 1860
aaaagtacaa gaagaaaaga ttaaaacctt acaaaaggaa agggaagata aggaagaaac 1920
cattgatata cttagaaaag aattaagcag aacagaacag ataagaaaag agttgagcat 1980
taaggcttcc tccctagagg ttcaaaaggc ttcaaatagaa ggtcgtttgg aagagaaaaga 2040
gtccttggtg aaacttcagc aagaggaatt gaacaaacac tcccacatga tagcaatgat 2100
ccacagttta agtggtggaa aaataaatcc agaaactgtg aatctcagta tatagacatt 2160
atggcatttt ggaatttgta atctcatgat atttttgatg tatttatcta ttggaggggg 2220
ggtgggtagg ggagtttaatt tgtgacttcg taacaataag aagttattat ctaatttagt 2280
aaagaccctg atctgttgca tgttttttat ttaatagttt gaatagaaat ttaattttct 2340
aagttttact tttgtttct ggcttttatg gcttaagggt ttctttgggt cttacattag 2400
aaaaatcattt ttaacctcca ttatcatttt tctaaagttc ttctcttttt cttagtgtgt 2460
ttctattctg ttttgctgt cttatttatt ctcatctgtg attatttaga tcttaagacc 2520
aaactttctt ggtataacag toctaaagat tacaataata aaatatagag agagtaaaag 2580
taaaaagtaa agtaaaaaga gagaaggtat gattttactg ctagagaagt ttgtctctga 2640
agaagcacaa aagaaaatat tagtgaattt aaaataattt ttatactgct gtagcataat 2700
ttctaaattt gaaaaaatg caatggtaat aaaatgtata aaaattagaa aactgtcatt 2760
gtgttaaact attacattta aatgattaca ttaaacacaa tagctgtctc ataaaaaatc 2820
taagaacttg tagaattatt tgtaaggat attagtgttt tttctctctt tttcaaatgt 2880
tgcatagtgg tccagatcat ctataattaa ttataagttt cttatgattt acataaggat 2940
aatttggatt ctactcttgc tgcataattca gaaaaatatt atttactgtt aattctgata 3000
tgaattgtat ttaatttggt gtactaatag tttttaattg gttcatttaa ttttaagca 3060
taggttaaat atctttattt tttttaaaat acaagggtcc ttctgtactt ttctattgaa 3120
tcatatcttc taactactct gtaagaaacg tttttacatc aatattttga ttgattttat 3180
tgctatacca atacagtcatt taattttact tgggtgtttc ttcacacata tgtctgtttg 3240
atttaagtg 3249

```

<210> 43

<211> 4133

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2738233CB1

<400> 43

```

gaaaaatgtg agaaaggaat agcagatttc ctggagaaac tacgaacttt caaaaagaag 60
ctttcgcagt ctctcccgga tcaccatgaa gagctccatg cagaacaaat gcgttgcaag 120
gaattagaaa atgcagtggg gatgacttga gatgacttga cccagttgag cctgctgaag 180
gacaccctct ctgcctatat cagtgtctgat gatatctcca ttcttaatga acgcgtagag 240
cttctgcaaa ggcagtggga agaactatgc caccagctct ccttaaggcg gcagcaata 300
ggtgaaagat tgaatgaatg ggcagtcttc agtgaaaaga acaaggaaact ctgtgagtgg 360
ttgactcaaa tggaaagcaa agtttctcag aatggagaca ttctcattga agaaatgata 420
gagaagctca agaaggatta tcaagaggaa attgctattg ctcaagagaa caaaatacag 480
ctccaacaaa tgggagaacg acttgctaaa gccagccatg aaagcaaagc atctgagatt 540
gaatacaagc tgggaaagggt caacgaccgg tggcagcatc tcctggacct cattgcagcc 600
aggggtgaaga agctgaagga gacctgggta gccgtgcagc agcttgataa gaacctgagc 660
agcctgagga cctggctcgc tcacatcgag ctcagagctgg ccaagccaat agtctacggg 720
tctgttaact cggaagaaat acagagaaag cttaatgagc agcaggagct tcagagagac 780
atagagaagc acagtacagg tgttgcatct gtctctcaacc tgtgtgaagt cctgctgcac 840
gactgtgacg cctgtgccac tgatgccgag tgtgactcta tacagcagggc tacgagaac 900
ctggaccggc ggtggagaaa catttgtgct atgtccatgg aaaggaggct gaaaaatcgaa 960

```

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| gagacgtggc | gattgtggca | gaaattttctg | gatgactatt | cacgttttga | agatttggtg | 1020 |
| aagtcttcag | aaaggacagc | tgcttttccc | agctcttctg | gggtgatcta | tacagtgtgc | 1080 |
| aagggaagaa | taaaagaaat | tgaggcttcc | cagcgacagg | tccacgagtg | cctgacgcag | 1140 |
| ctggaactga | tcaacaagca | gtaccgccc | ctggccagg | agaaccgcac | tgattcagca | 1200 |
| tgtagcctca | aacagatggg | tcacgaaggc | aaccagagat | gggacaacct | gcaaaagcgt | 1260 |
| gtcacctcca | tcttgccgag | actcaagcat | tttattggcc | agcgtgagga | gtttgagact | 1320 |
| gcgcggggaca | gcattctggg | ctggctcaca | gagatggatc | tgcatctcac | taatattgaa | 1380 |
| catttttctg | agtgtgatgt | tcaagctaaa | ataaagcaac | tcaaggcctt | ccagcaggaa | 1440 |
| atttcaactga | accacaataa | gattgagcag | ataattgccc | aaggagaaca | gctgatagaa | 1500 |
| aagagtggag | ccttggtatg | agcgatcatc | gaggaggaa | tagatgagct | ccgacgggtac | 1560 |
| tgccaggagg | tcttcgggcg | tggtgaaaga | taccataaga | aactgatccg | cctgcctctc | 1620 |
| ccagacgatg | agcagacact | ctcagacagg | gagctggagc | tggaagactc | tgcatctctg | 1680 |
| tcggacctgc | actggcacga | ccgctctgca | gacagcctgc | tttctccaca | gccttccctc | 1740 |
| aatctctccc | tctcgtctgc | tcagccccct | cggagcagag | ggtcaggacg | agacaccccc | 1800 |
| gctagtgtgg | actccatccc | cctggagtgg | gatcacgact | atgacctcag | tcgggagctg | 1860 |
| tgctctggca | tgccagagc | tctgccctct | gaggatgaag | aaggctcagga | tgacaaagat | 1920 |
| ttctacctcc | ggggagctgt | tggttatca | ggggaccaca | gtgccctaga | gtcacagatc | 1980 |
| cgacaactgg | gcaaaagccct | ggatgatagc | cgctttcaga | tacagcaaac | cgaataatc | 2040 |
| attcgcagca | aaactcccac | ggggccggag | ctagacacca | gctacaaagg | ctacatgaaa | 2100 |
| ctgctgggag | aatgcagtag | cagtatagac | tccgtgaaga | gactggagca | caaatgaag | 2160 |
| gaggaagagg | agagccttcc | tggtttgtt | aaactgcata | gtaccgaaac | ccaaacggct | 2220 |
| gggtgtgattg | accgatggga | gcttctccag | gcccaggcat | tgagcaagga | gttgaggatg | 2280 |
| aagcagaaac | tccagaagtg | gcagcagttt | aactcagact | tgaacagcat | ctgggccttg | 2340 |
| ctgggggaca | cggaggagga | gttggaaacg | ctccagcgtc | tggaactcag | cactgacatc | 2400 |
| cagaccatcg | agctccagat | caaaaagctc | aaggagctcc | agaaagctgt | ggaccacccg | 2460 |
| aaagccatca | tctctccat | caatctctgc | agccctgagt | tcacccaggc | tgacagcaag | 2520 |
| gagagccggg | acctgcagga | tcgcttgcg | cagatgaatg | ggcgctggga | ccgagtgtgc | 2580 |
| tctctgctgg | aggagtggcg | gggctgctg | caggatgccc | tgatgcagtg | ccaggggttc | 2640 |
| catgaaatga | ggcatggtt | gttctttatg | ctggagaaca | ttgacagaag | gaaaaatgaa | 2700 |
| attgtcccta | ttgattctaa | ccttgatgca | gagatacttc | aggaccatca | caaacagctt | 2760 |
| atgcaataaa | agcatgagct | gttggaaatc | caactcagag | tagcctcttt | gcaagacatg | 2820 |
| tcttgccaac | tactggtgaa | tgctgaagga | acagactggt | tagaagccaa | agaaaaagtc | 2880 |
| catgttattg | gaaatcggt | caaactctc | ttgaaggagg | tcagtcgtca | tatcaaggaa | 2940 |
| ctggagaagt | tattagacgt | gtcaagtagt | cagcaggatt | tgtcttctctg | gtcttctgct | 3000 |
| gatgaactgg | acacctcagg | gtctgtgagt | cccacatcag | gaaggagcac | cccaaacaga | 3060 |
| cagaaaacgc | cacgaggcaa | gtgtagtctc | tcacagcctg | gacctctgt | cagcagtcga | 3120 |
| catagcaggt | ccacaaaagg | tggtctcgat | tcttcccttt | ctgagccagg | gccaggctgg | 3180 |
| tcgggcccgcg | gcttctctgt | cagagtctct | cgagcagctc | ttccccctca | gcttctctctg | 3240 |
| ctctcctcca | tcgggcttgc | ctgccttgta | ccaatgtcag | aggaagacta | cagctgtgcc | 3300 |
| ctctccaaca | actttgccc | gtcattccac | cccattgctca | gatacacgaa | tgccctctct | 3360 |
| ccactctgaa | ctaagcagat | gccatctgca | gaagtgtctg | tagcataagg | aggatcgggt | 3420 |
| cataagcaat | cccaactac | caacaagagg | accttgatct | tggcgaaagc | cctcgggtgtg | 3480 |
| gcagctttag | ccctcctcca | gatcacatgt | gtgcaaatga | tggtctcaga | ggtggaagat | 3540 |
| aaacagtgc | aggggaacaa | acagacaaca | agaaggtttg | gaagaaatct | ggtttgagac | 3600 |
| tctgaacctt | agcactaagg | agattgagta | aggacctcca | aagttccccg | gactcatgaa | 3660 |
| ttctggggccc | ttggccatt | ctgtgcacag | ccaaggactt | cagtagacca | tctgggcagc | 3720 |
| tttcccatgg | tgctgtctca | accatcagat | aaatgacct | cccaagcacc | atgtcagtgt | 3780 |
| cgtacaatct | accaaccaac | cagtgtgtaa | gagatttttag | aaccttgtaa | catacaattt | 3840 |
| ttaagagctt | atatggcagc | ttccttttta | ccttggtttt | ctttggggca | tgatgtttta | 3900 |
| acctttgctt | tagaagcaca | agctgtataat | cttaaaaggca | ctttttttta | gaggtataaa | 3960 |
| gaaaaactag | atgtaataaa | taagatcatg | gaaggcttta | tgtgaaaaaa | gttgaatgtt | 4020 |
| atagtaaaaa | aaaaaaagat | atztatgtat | gtacagtttg | ctaaagccaa | gttttgtttg | 4080 |
| tattgatttc | tttgcattta | ttatagatat | tataaaataa | aaaaaaaaaa | aaa | 4133 |

<210> 44

<211> 1754

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1833116CB1

<400> 44

| | | | | | | |
|------------|------------|-------------|------------|------------|------------|-----|
| cggaaggga | aaacaactac | ggctgcgggtg | tggttggttg | tgagatgacg | accttagtgc | 60 |
| tggaataatg | agcttacaac | gccaaaatcg | gttacagcca | tgaaaatgtg | tcggttattc | 120 |
| ctaattgtca | gttccggtca | aaaacagcac | gtcttaaaac | ttttactgcc | aaccagatag | 180 |

```

atgaaataaa agacccttct ggactctttt acatcctccc ttttcaaaag ggctacttgg 240
tgaattggga tgttcagaga caagtttggg attacctttt tggaaaagaa atgtatcagg 300
ttgatttttt agatactaatt attattatca ctgaaccata ctttaacttc acttcaattc 360
aagaatcaat gaatgaaatt ctatttgaag aataccagtt tcaagcagta ttaagagtaa 420
atgctggggc tctcagtgca cataggtatt tccgagataa tccttccgaa ttatgctgta 480
tcattgttga tagtgatat tcctttacac atatagttcc ttattgtaga agtaaaaaga 540
aaaaagaagc aattattcgg ataatgtgg gaggaaaact cttaaccaat catctaaagg 600
agatcatatc ttacaggcag ctacatgtta tggatgaaac acatgtgatt aatcaagtga 660
aagaagatgt atgctatgtg tctcaggatt tttatagaga catggatatt gcaaagtgtga 720
aaggagaaga aaatacagta atgatagact atgtcttggc tgacttcagt acaattaaaa 780
agggcttttg taagccaagg gaagagatgg tgttgagttg aaaatacaaa tctggggaac 840
aaattcttcg tttggccaat gagagatttg ctgttccgga aatactcttt aatccttctg 900
atataggcat tcaagaaatg ggaattccag aagctattgt ctattcaatt caaaatctac 960
ctgaagaaat gcagccgcat ttttttaaga acattgtctt gacaggagga aattcccttt 1020
tcccaggaatt tagggatcgg gtttactcag aagttcagatg tcttactcca acagattatg 1080
atgtttctgt tgtgctgcct gaaaacccta ttacttatgc ctgggaaggt ggaaaattga 1140
tatcagagaa tgatgatttt gaagatatgg tggtaacaag agaagattac gaagaaaatg 1200
gacatagcgt ctgtgaagag aaatttgata ttttaagcaac atttttgaat gaaagtgtgtg 1260
accataaggg taattttcaa agttcctttt aaaagaggtt aaggaaactgt gttacctttt 1320
gtcctaagaa aaaggcttga atttatgtaa atactttgat cgattgctaa ttttcaagg 1380
cttcttaggt aggttactac agtaaaactgt aactcagtc acattttcat ttaggagcta 1440
gactaccata acaatgctta tgctgtttcc aagggtaggt tatttttcat taaaagaaga 1500
atgaatgcat ttttaagttta attcttcata gctgaaagca caaatttaac ggcttctactg 1560
gacagttttc cttggaaggt agttttgtgt gactgtgact aaactatttt atttttaaat 1620
gtcattctta tttatacatt ctaaagttgg aaagactgat cttatatgtg tataatgttt 1680
attttgtacc tagagtacat ttaaaagggt ggagactaag ctaataaagt ttttttggcc 1740
actaaaaaaa aaaa

```

<210> 45
<211> 2713
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 001799CB1

```

<400> 45
ttgaatggcc ctgagtggga ctcgccccag aagccgaggg actctctagg ctgccgggcg 60
ctggctcgta gcgcgcaggc tgggctgagg cgccgcggta ccatgaggcg ccggtactta 120
agagattatg gcatcagaaa ccacaaatgt taaaaaacgg aacttttgta ataagattga 180
ggatcatttc attgatcttc ctagaaaaaa gatctctaatt ttcactaata agaacatgaa 240
ggaggttaag aaatctccaa aacagttggc tgcttacata aatagaacag ttggacaac 300
tgtgaaaagc ccagataaac ttcgtaaagt gatctatcgc agaaagaaag ttcacatccc 360
ctttccaaat ccttggttaca gaaaaaaaca gtcccctgga agtgggggct gtgacatggc 420
aaataaagaa aatgaactgg cttgtgcagg ccacctgcct gaaaaattac accatgatag 480
tcgaacatat ttggttaact ccagtgatcc tggttcttca gagacagaaa gcccatcatc 540
aaaatatagt ggggtttttt ctgaggtttc tcaggaccat gaaacaatgg cccaagtttt 600
gttcagcagg aatatgagat tgaatgtagc tttaactttc tggagaaaga gaagtaaaag 660
tgaacttgta gcttatttgt tgaggataga agatcttggc gttgtggtag attgccttcc 720
tgtgctcacc aattgtttac aggaagaaaa acaatatatc tcacttggct gctgtgttga 780
cttgttgccct ctagttaaag cactacttaa aagcaaattt gaagaatatg ttatagttgg 840
tttaaaactgg cttcaagcag tcattaaaag gtggtggtca gaactatcat ccaaaacaga 900
aattataaat gatggaata ttcaaatttt ttaagtggat tatgggaaca 960
ggaaaaccat cttactttgg ttccaggata tactggtaat atagctaagg atgtagatgc 1020
ttatttatta cagttacatt gagagatttc atctactaaa gagcatttgg tttttcaaaa 1080
catccctgaa ctgtataatt tacaaaaaaa aaagtctcgt ctgagaactg tgaactgtgg 1140
aagaaatcaa aactattttt tcttttataaa agccacgtaa tgaaccact aatgaaatcc 1200
cagcaatctg cttcacattg aagtggaaaa atatccaaaa ggagcagctt caatttcatt 1260
gaggtgaaag tgcactatga agattgttca cctttgctgc atttgggagt tatatggtta 1320
tttggttaaca ttaagaacta ctggatttta atgcaatcct gcataaaaaa ataatttata 1380
ctatgtgaaa aaataagaca ggacttacca ctaggaacca ccaagacca tcatcattaa 1440
cttttttaag attgtgtttt attaaaaaaa aaaaacactt aaatgtgtgc agctattttc 1500
ttatgttgaa aagactgaaa gtttaaaaaca tgaaaaaaat caatatataa cattttttgt 1560
tcacactgag atactgtgta tgtaaaatgc ctaaatattt aataagccaa tgtgttatga 1620
taccaatatc tgttttataa aactaaaacc aaccatgctt ctggcatgat aaaatcatgg 1680
aattaaatca ggggtttaca ttcttgtaga gtgttcttga aacactctct gcaccatttt 1740

```

```

taaaacttga gaatagtttt agtatctctg atattttttg ccagaatcat catgtcatgt 1800
atgaatgtgt tatccctatc taaggaaaaa ggtgaatatg tttttgtatg aatgtttaac 1860
tggaatgtgc catggacttg gctaatttat atttactttt tattgtacat agatttctaa 1920
tatttttcat tcctgtatca tttaaacttc ctccatttga gtaaattcac taaatatttc 1980
tattttttgc ttttttaa atctgatttta tatgaattct aattcttttt cactacatat 2040
gttttaaaga gttacatata gtgatttaga atggtttaca gttaatgctg atctgtatt 2100
ttaaatcca acactttgtg tcactacctc ctctaattgt tagtatgata tgctagcaga 2160
ctgtatgagg tcttttttta aaataccact tttagtgtca gtgaacaaa ttctggaatg 2220
tcttaacagc tctaaatctt acttgtcttg aaaatgattg ggttttaata ccactgctgg 2280
tggttcacac atcatcccat ccttaatatg cctgacaggc atctgagcaa aggttttag 2340
taattgaatt tctctgcagt agtcttcaa gcacttgaat gtaaaccttt agcatttatt 2400
cgtttaatga ctactgatac gaatctcaag cagatttctt gctcttaaaa gttatgtttc 2460
actgagttct ggttttgtgt agctatattt tatatagcta gatattcttc acagtgaaca 2520
tgaattgtaa taattgggta tttccttaag tcttttagatt ataataattt cagattattg 2580
cacgttgtag atttgagagg tgagttattt aagaggccag ttttcaggac atgggaattt 2640
gaattgttaa cctgttatct ctgtgaaact tttaacatga taaaatataa cttttctttg 2700
tgcttaaaaa aaa 2713

```

<210> 46
 <211> 1768
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 119814CB1

```

<400> 46
ttttttttt ttttttttg ctcttttagaa gaggttatat ttttattatc cttatttttg 60
agaacttttc cttataaaat tttttttcca gatttcttat gaactcaagt tagtggttaa 120
gctttggatt ccactgttaa cagttttatgt aaaaacactt aacaaattgc catttatatg 180
ccaaactata gctcaagaac actctgtttt agaaaaatta cgcattagat caggaagcct 240
catatatatg tgccctctgg acttcatttg agtcacatt tagccagaaa agcaatgact 300
tctatatctc ttatggaaac caatgtaaca taaattaatg ttctaaatat agaaattaag 360
agttcataaa gagactgagg ttgcatgtaa aagagttatg gtttgagaca gtctaaaaat 420
actatgttaa tttcaaggat cttatttcca atgttttgtt taaaaaatta taaatacttt 480
tgagctcttg ctttgcattt caatcgcaa cccactcaga tacgggaact gtttaaattc 540
atatatggac aaataggttt cagtgtatgca atactttaa attctgccat ctcttgtgt 600
ttttctttct aggtgagtgg actgccagct cctgatgtgt catggtatct aatggaaga 660
acagttcaat cagatgattt gcacaaaatg atagtgtctg agaaggtct tcattcactc 720
atctttgaag tagtcagagc ttcagatgca ggggcttatg catgtgttgc caagaataga 780
gcaggagaag ccaccttcac tgtgcagctg gatgtccttg caaagaaca taaaagagca 840
ccaatgttta tctacaaacc acagagcaaa aaagtttttag agggagattc agtgaaacta 900
gaatgccaga tctcggtat acctccacca aagcttttct ggaaaagaaa taatgaaatg 960
gtacaattca acactgaccg aataagctta tatcaagata acactggaag agttacttta 1020
ctgataaaaag atgtaaaaaa gaaagatgct ggtgtgtata ctgtgtcagc agttaatgaa 1080
gctggagtga ctacatgtaa cacaagatta gacgttacgg cacgtccaaa ccaaactctt 1140
ccagctccta agcagttacg ggttcgacca acattcagca aatatttagc acttaatggg 1200
aaaggtttga atgtaaaaaa agcttttaac ccagaaggag aatttcagcg tttggcagct 1260
caatctggac tctatgaaag tgaagaactt taataacttt accaacattg gaaaacagcc 1320
aactacacca ttagtaatat atttgattac atttttttga aattaatcca tagctgtatt 1380
aacagattat ggttttaatt aggtaatat atttataata atttataata ttatttatcc 1440
tttgactctt gcacattcta tgtaccctc cgatttgtga agcctacagg aaatctgggt 1500
atatggattt gtaactgcag aagactatct taaaatacag gatttttaaca ttttagtcat 1560
gcacatttaa caattacagg ttataaatta gtatcaactt tttaaacaca tctaattgct 1620
gtaataacgt ttactggtac tgctttctaa atactgtttt acccgttttc tctttagtag 1680
atactaacat ggtatagatt atctgagtgt tccacagttg tatgtcaaaa gaaaataaaa 1740
ttcaaatatt taaaacggaa aaaaaaaa 1768

```

<210> 47
 <211> 3287
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1295420CB1

<400> 47

```
cgccccctga caccgacacc ccattgctcc cacagtctcc ccagtctcca ctttggtccc 60
cagcgctgtc tgccccagga tttgcctgaa ggctgcccc aactctgcac ccgcccccg 120
agggccaccg aggaccatga ctaagacaga tcttgcccc atggccccgc caccocgagg 180
agaggaggaa gaagaggagg aggaggatga acccgtcccc gagggcccca gccccacca 240
ggagcgccgg cagaagcctg ttgtgcacc ctccgaccc gccccctcc ctaaggacta 300
cgctttcacc ttctcgatc ccaatgaccc ggctgcccag gagatcctgt ttgacctca 360
gaccaccatc cccgagctgt ttgccattgt gcgccagtgg gtgccccaa gtcagacaaa 420
gatagacgtc atcggcaatg agattctgct cccactatgc gtgcaaagct gggggccacg gagtccggga 540
gctgaccgac atgacactgc tccactatgc gctgctggcg ctggggcgag atgtgacgt 600
cccccgcca gccgtgcgcc tctcgacga tcaactacgc gcctattttg atgtgcccga 660
gcgcagccgc tggaccaaca tgaacgcgt tccactacgc gtgaactcca cgtgcagtga 720
cctcgtgcgt gtgctgctga aggggtgcag gccgcgagtg gtgaactcca cgtgcagtga 720
cttcaaccac ggctcagccc tgcacatcgc tgcctccagc ctgtgcctgg gcgccgcaa 780
atgtttgtgc gagcacggcg ccaaccctgc gctgaggaat cgaaggagac aggtgcggcg 840
ggaggtgggt ccagatccta tggacatgct cctggacaag gcagaggcgg cactggtggc 900
caaggagctg cggacgcttc tgggaagagg agtgccacta tcttgcccc tccccagggt 960
cacgtacccc aactatgaca acgtcccagg caatctcatg cttagcgcac tgggcttgct 1020
cctgggagac cgcgtgctgc tggatggcca gaagacgggc acactgcggt tctgtgggac 1080
cacggagttt gccagcgccc agtgggtggg cgtggagctg gacgaacctg agggcaagaa 1140
cgatggcagc gttggggcg ttcggtaact catctgccc cccaagcagg gtctctttgc 1200
ctcgtgtccc aagatctcca aggcagtggc cgcaccccc tctctgtca cctccacacc 1260
ccggaccccc cggatggact tctccctgt tcccgcaaa ggccgcaggg aacacaaagg 1320
caagaagaag accccatcat cccactctct gggcgagctt cagcagcgtg acggggccaa 1380
ggctgaggtt ggagaccagg tccctgtcgc gggccagaag caggggatcg tgcgcttcta 1440
cggaagaca gactttgccc caggttaact gtatggcatt gagctggacc agcccacagg 1500
caagcatgat ggctctgtct tcgggtgtcc gtacttact tgccccccga ggcatgggg 1560
cttcgacca gcatccgta ttcagagat ttggcgatcc actgattccc ccggggacag 1620
cgttgagacc aaaaaagtgc atcaagtac aatgacgcag cccaaacgca ccttcaccac 1680
agtcggacc ccaaaggaca ttgcatcaga gaactccatt tccaggttgc tgttctgctg 1740
ctggttcccc tggatgctga gggcggagat gcagtcttag agggccctga accctgagtc 1800
agagacagag tccccactag catctcctga cacccgagga gccctgagtc accctgagat 1860
agagattccc gtaaacacat ccagagtga ccccctgtt agccagccct cgtacattga 1920
ggccccatta ttaacagata ctcccataat aacccccaaa tacagacccc atgtcaccca 1980
gaaagagatt ccctgagtag caccttcagg ctagtcccta tccccacccc ctcagagcag 2040
attcccagat taacagattt ccataatcacc ccaaatgatg gtgacctct ccacataatg 2100
cattacaaca gaacattctt gaatcaccca accctggatc agaaaacctcc ccattaacaa 2160
aactgcccc ttaagtcttc ttgaaataaa cataggtcac acccccaaag caaaagagta 2220
acagacattc atgtcattgt tccccattta acatcagtec totcaagatg tctcaagatg 2280
atggtcacc tgaagccctt agattccaac cctcaatca gagacttctt tcattaacaa 2340
agacccttgt tcttatcctt caagaagaaa cccaccataa ccagccact gtcacccta 2400
atttacagac accaaaacag tccctggaag gctaattaca ggacccccca agtcttcta 2460
ccctctgcac cctcaagaaa cccccagtc cttgtatgaa gccaccccca catggccac 2520
agctcctgtg ctggccagac tcccagaaaa ttctctatth tttaagtaac gacttcccc 2580
tttgggggac cccaaaattt ggaggcccca tcttaggact ctggggatcc caaacctag 2640
agtacacacg tcccaaaact cctgtgccc tcaagtccta cagccctag aagaccccaa 2700
tgccgtaact ctaggacccc ccaaatcatg gaatcccaaa tccccaggga atcccaaatt 2760
tgaaaatcca atcccaagtc cccaggaaac ccaatcatga ggtccttgtg cctggtatgg 2820
aggagactgc agtcaggata tgcattccag gctcccagac acctcaagcc ctattcacag 2880
gcaccaggaa accccacaca ggaattccca tccctggaaa ctggagaatt tcaatgcccc 2940
gagtccatgg gtttcaagac accaaatttc aagagcccca gccctaagg gaccccaaat 3000
cctaaagcct ccatctctaa taaatggaag gccccaaggc cctgagggga tctcaaatcc 3060
tggaaccccc atttcaatct acgttctagt cactggcctc aaaggacccc acagcacctg 3120
ggccagacca acagctcgag ggagaacctg aaggccagg ggttcagg gggacctggg 3180
gccccgacca ccaaggacag ctacgactg ccccttact gcatgtcccc aaactcagca 3240
tgactcctgt cctcttcaat aaagacgttt ctatggccaa aaaaaaa 3287
```

<210> 48

<211> 1748

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1309364CB1

<400> 48

```

cgccatgttt gaaaagtgat gacgggttgac gtttgctgat ttttgacttt gcttgtagct 60
gtccccgaa ctccgctgtct tccgttcggc ggccggcact gtaggtgagc gcgagaggac 120
ggaggaagga agcctgcaga cagacgcctt ctccatccca aggcgcgggc aggtgccggg 180
acgttggggc tggcgggtgtt ttctgtcgtgc tcagcgggtgg gaggaggcgg aagaaaccag 240
agcctggggag attaacagga aacttccaag atggaaactt tgtctttccc cagatataat 300
gtagctgaga ttgtgatca tattcgcaat aagatcttaa caggagctga tggtaaaaaa 360
ctaccaaga atgatcttta tccaaatcca aagcctgaag tcttgacat gatctacatg 420
agagccttac aaatagtata tggaaatcga ctggaacatt ttacatgat gccagtgaac 480
tctgaagtca tgtatccaca tttaatggaa ggcttcttac cattcagcaa tttagt tact 540
catctggact cttttttgcc tatctgcgg gtgaatgact ttgagactgc tgatattcta 600
tgtccaaaag caaaacggac aagtcgggtt ttaagtggca ttatcaactt tattcaactc 660
agagaagcat gccgtgaaac gtatatggaa tttctttggc aatataaatc ctctgcggac 720
aaaaatgcaac agttaaacgc cgcacaccag gaggcattaa tgaaactgga gagacttgat 780
tctgttccag ttgaagagca agaagagttc aagcagcttt cagatgggat tcaggagcta 840
caacaatcac taaatcagga ttttcatcaa aaaacgatag tgctgcaaga gggaaatccc 900
caaaagaagt caaatatttc agagaaaacc aagcgtttga atgaactaaa attgttggtg 960
gtttctttga aagaaatata agagagtttg aaaacaaaaa ttgtggattc tccagagaag 1020
ttaagaatt ataaagaaaa aatgaaagat acggtccaga agcttaaaaa tgccagacaa 1080
gaagtgggtg agaaatatga aatctatgga gactcagttg actgcctgcc ttcatgtcag 1140
ttggaagtgc agttatatca aaagaaaaata caggaccttt cagataatag ggaaaaatta 1200
gccagtatct taaaggagag cctgaacttg gaggaccaa ttgagagtga tgagtcagaa 1260
ctgaagaaat tgaagactga agaaaattcg ttcaaaagac tgatgattgt gaagaaggaa 1320
aaacttgcca cgcacaatt caaaaataat aagaagcatg aagatgttaa gcaatacaaa 1380
cgacagtaaa ttgaggttg caataaagtt caagaaaaaa gaggtgctgt ctatgaacga 1440
gtaaccacaa ttaatcaaga aatccaaaaa attaaacttg gaattcaaca actaaaagat 1500
gtctgtgaaa gggagaaact gaagtcccag gaaatatttc taaacttgaa aactgctttg 1560
gagaaatacc acgacggtat tgaaaaggca gcagaggact cctatgctaa gatagatgag 1620
aagacagctg aactgaagag gaagatgttc aaaatgtcaa cctgattaac aaaattacat 1680
gtctttttgt aaatggcttg ccatctttta attttctatt tagaaagaaa agttgaagcg 1740
aatggaag 1748

```

<210> 49

<211> 2163

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1315267CB1

<400> 49

```

gatgaattga gggaacagcc ttgtaaaatc aggaaagccg tccaaaagag cacttctgaa 60
aatcagactg aatggaatgc acgtgacgat gaaggtgttc caaatagtga cagtagcact 120
gactctgagg aacagcttga tgttaccata aaaccatcga ctgaggatag agagaggggc 180
atcagcaga gagagatag ccacaaagtc tgtgatgata aggggccttt taaggacacc 240
aggacccaag agataaaaag gagagatgtt gatctggatt tgtctgataa agattacagt 300
agcgatgagt ctatcatgga aagcataaaa cataaagtgt ctgagccctc gagatcctca 360
tccctaagtc tgagtaaaat ggactttgat gatgaaagaa cttggactga ccttgaagag 420
aatttgtgta accatgatgt tgttcttggg aatgaatcca cttatgggac gccgcagaca 480
tgctacccta ataataaagt aggtatcctg gacaaaacaa taaaaggaa gattgcacca 540
gtcaagaggg gagaagactt gagcaagtcc aggaggagca gaagtcctcc tacatcggag 600
ctgatgatga aattctttcc ttctttgaaa ccaaaaccaa agtcagattc acacttggga 660
aatgaactca agttaaacat aagtcaagac caaccacctg gtgacaatgc tcgatcccag 720
gttttgagag agaaaattat tgaattggaa acagaaatag aaaagtttaa agctgagaac 780
gcatcttttag ctaaacttcg cattgaacga gaaagtgcct tggaaaaact caggaaagaa 840
attgcagact tcgaacaaca gaaagcaaaa gaattagctc gaatagaaga gtttaaaaag 900
gaggagatga ggaagctaca aaaggaacgt aaagtttttg aaaagtatac tacagctgca 960
agaacttttc cagataaaaa ggaacgtgaa gaaatacaga ctttaaaaca gcaaatagca 1020
gatttacggg aagatttgaa aagaaaggaa accaaatggt caagtacaca cagccgtctc 1080
agaagccaga tacaatgtt agtcagagag aacacagacc tccgggaaga aataaaagtg 1140
atggaaagat tccgactgga tgcctggaag agagcagaag ccatagagag cagcctcgag 1200
gtggagaaga aggacaagct tgccaacaca tctgttcgat ttcaaaacag tcagatttct 1260
tcaggaaacc aggtagaaaa atacaagaaa aattatcttc caatgcaagg caatccacct 1320
cgaagatcca agtctgcacc tctctgtgat ttaggcaatt tggataaggg acaagctgcc 1380
tctcccaggg agccacttga accactgaac ttcccagatc ctgaatataa agaggaggag 1440
gaagaccaag acatacaggg agaaatcagt catcctgatg gaaaggtgga aaaggtttat 1500
aagaatgggt gccgtgttat actgtttccc aatggaactc gaaaggaagt gagtgcagat 1560

```



```

gggaagacca tcaactgtcac tttctttaat ggtgacgtga agcagggtcat gccagaccaa 1620
agagtgtatct actactatgc agctgcccag accactcaca cgacataccc ggaggggactg 1680
gaagtcttac atttctcaag tggacaaata gaaaaacatt acccagatgg aagaaaagaa 1740
atcacgtttc ctgaccagac tgttaaaaac ttatttccctg atggacaaga agaaagcatt 1800
ttcccagatg gtacaattgt cagagtacaa cgtgatggca acaaaactcat agagtttaat 1860
aatggccaaa gagaactaca tactgcccag ttcaagagac gggaataccc agatggcact 1920
gttaaaacccg tataatgcaaa cgggtcatcaa gaaacgaagt acagatccgg tcggataaga 1980
gttaaggaca aggagggtaa tgtgctaattg gacacggagc tgtgacgatc ctcatgtgat 2040
catgaagtaa cagtaactga ctttttatgt taaaaaatgt acatttactg tggattctgt 2100
ttaatttatt gtgtatgtgt ggggaaaaaga ttggattcta aaataaaagt ttaccctgtg 2160
gca 2163

```

<210> 50

<211> 1615

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1403289CB1

<400> 50

```

gtggttcaga ggcagcttct agacctgcag gagggagatt gtattcagag gaagagcatc 60
atthttggcaa catctgaaag tgaanaacgga agccagaaac acttggccag ccctggggga 120
tttttttctt ctatgcctct ctggtggaat gacatttgct gtgtaggcat ctttcctctg 180
actgtatttc ttggccttga agagtactga gtttaaaaag acagtatgtg acagtccatg 240
gaaattgcct cttctgtgaa atctcgccac ctgctccgaa gacatgttgt tgtctcccaa 300
attctcctta tccaccattc acatacgact gacggccaaa ggattgcttc gaaaccttcg 360
acttccttca ggggttagga gaagcactgt tgttttcac acagttgaaa agagcaggca 420
aaagaatcct cgaagcttat gtatccagcc acagacagct cccgatgcgc tgccccctga 480
gaaaacactt gaattgacgc aatataaaac aaaatgtgaa aaccaaagtg gatttatcct 540
gcagctcaag cagcttcttg cctgtggtaa taccaagttt gaggcattga cagttgtgat 600
tcagcacctg ctgtctgagc gggagggaagc actgaaacaa cacaaaacc tctctcaaga 660
acttggttaac ctccggggag agctagtac tgcttcaacc acctgtgaga aattagaaaa 720
agccaggaat gagttaaaaa cagtgtatga agcattcgtc cagcagcacc aggrtgaaaa 780
aacagaacga gagaatcggc ttaaagagtt ttaccagg gagtatgaaa agcttcggga 840
cacttacatt gaagaagcag agaagtacaa aatgcaattg caagagcagt ttgacaactt 900
aatgctgcg catgaaacct ctaagttaga aattgaagct agccactcag agaaaactga 960
attgctaaag aaggcctatg aagcctccct ttcagaaatt aagaaaggcc atgaaataga 1020
aaagaaatcg ctggaagatt tactttctga gaagcaggaa tcgctagaga agcaaatcaa 1080
tgatctgaag agtgaaaatg atgctttaa tgaaaaattg aaatcagaag acaaaaaaag 1140
aagagcaaga gaaaaagcaa atttgaaaaa tctcagatc atgtatctag aacaggagtt 1200
agaaagcctg aaagctgtgt tagagatcaa gaatgagaaa ctgcatcaac aggacatcaa 1260
gttaatgaaa atggagaaac tgggtggaca caacacagca ttggttgaca aattgaagcg 1320
ttctcagcag acggagcagg ctgttctgca agagtcgctg gagaaggagt cgaaagtcaa 1380
gcagctttcc acggagcagg ctgttctgca agagtcgctg gagaaggagt cgaaagtcaa 1440
caagcgactc tctatggaaa acgaggagct tctgtggaaa ctgcacaatg gggacctgtg 1500
tagccccaag agatcccca catcctccgc catcctttt cagtcaccaa ggaattcggg 1560
ctccttacat agccccagca tttcaccagc atgacacgtc cccaaagtcc acaga 1615

```

<210> 51

<211> 1356

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1607607CB1

<400> 51

```

ccaccccggg aacccggaag tggaggagga ggcgcgcgcg cgcgcgcgcg ggcggtgcg 60
gtggccaagc aggcagatag tgcctgacct gttcccgga gcgtgtctgg gtttgggggc 120
gggagacagg ctgagccgcc tggcgccct ggctgtacg gggcggggga ggccatggcc 180
tcggctgagt tgcaggggaa gtaccagaag ctggctcagg agtactcgaa gcttcgggct 240
cagaatcagg ttctgaaaaa aggtgttgtg gatgaacaag caaattctgc agctttaaag 300
gagcaactga aatgaagga tcagtcattg agaaaactac aacaggaaat ggacagtttg 360
acatttcgaa atctgcagct tgccaagagg gtgaactac ttcaagatga actagctcta 420

```

```

agtgaaccac gaggcaagaa aaacaagaaa agtggagaat cttcttctca gttgagtcaa 480
gagcagaaga gtgtctttga tgaagatctg caaaagaaga tagaagagaa tgaacggttg 540
catatacaat tttttgaagc tgatgagcag cacaagcatg tggaaagcaga gctgaggagt 600
cgactggcca ctctgggagac agaagcagcc cagcaccaag ctgtggttga cggctctacc 660
cggaagtaca tggaaacccat tgagaagctg cagaacgaca aggtctaaact agaagtgaaa 720
tctcagactc tagaaaagga agccaaggaa tgtcgacttc gaacggaaga atgtcaatta 780
cagttaaaga ctcttcatga agatttgtca ggtagattag aggaatcctt atcaatcatc 840
aatgaaaaag taccttttaa tgatacaaaa tatagtcagt acaacgctct gaacgttcca 900
ctccacaata ggagacacca gctgaagatg tttcatacct acacagaaca gaggattcaa 1020
gttcaggatc ttgtgacggc tcttctaaac cattgacact atatctccat tgaatcagaa gttctcaca 1080
atcttctctg tttgtctgac ctatgtccgc cctcttgagg aaggaaatgct tcatttattt 1140
taccttcatg aaaatgcgtc tgtgactgtc ttggagacaa ctgtgaaatt gaaaactttt 1200
gaaagtatca ctgaggatac tgtagactgtc ctaggaaga ttcttcccta tcagttaaaa 1260
tcagaacact taacctccta catatgtttt cagggaactt ggggaattgt ggtgtgaacc tgatctggct 1320
aggtagttac ccccgaggc atgtgaaaga aaaaaa
ggcgtaata aaatattaac

```

<210> 52

<211> 1268

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1660025CB1

<400> 52

```

gatggaagta agccttgagt attgagggtt tcagaggggg tgcattctacc ccagccccag 60
ggagtgggga ggcggaagga ggacctgcgg caggccctct tcggcagctc ctccggccccg 120
gtttccctcg gcgtgctact gtgcgctcga tccagcacca tggggaagcg ggacaatcgg 180
gtggcctata tgaacccaat agcaatggcg agatcaaggg gtccaatcca gtcttcaggg 240
ccaacaatac aggattatct gaatcgacca aggcctacct gggaagaagt aaaagagcaa 300
ctagaaaaaaga aagaagaagg ctccaaggct ttggctgaat ttgaagaaaa aatgaatgag 360
aactggaaga aagaactgga aaaacacagg gagaaattgt taagtggaag tgagagctca 420
tccaaaaaaa gacagagaaa gaaaaaagaa aagaagaaat ctggtaggta ttcattctct 480
tcttcatcaa gctctgattc ttccagcagt tcttctgatt ctgaagatga ggataagaaa 540
caaggaaaaac ggagaaagaa aaagaagaac cgttcacata aatcttctga aagctccatg 600
tcagaaactg aatcagacag taaggatagt ttaaaaaaga aaaagaagtc aaaagatgga 660
actgagaaag aaaaggatat taaaggactc agcaaaaaaga gaaagatgta ttctgaagat 720
aaacctttat catctgagtc cttgtcagaa tcagagtata ttgaggagggt gcgagcaaaa 780
aagaagaaaaa acagtaagaa gaagaaaaag aaggctgcta gttcaagtcc tgactacca 900
cataagaaac aaaaatcagg attcccttat aaagaagtg caatgtctga ggaaatttca 960
taacattaag ctacaacata tttactaaaa tgcatgaatt ttcttgtttt tagaattatt 1020
cctggactat tcagtagcca ctcatgtgaa agggccataa atgttgccctg 1080
ctgcttgaac atctattttt ttctcttcca gtgcttgata actctgggag ataatacact 1140
gcagtcgtac tagtggttaa gatatttggg aataaaatta atacttttga ctagaagcgt 1200
ctaaggataa accaacagaa attgaatctg gatacatctt taagatgtaa tcagaaatga 1260
ccagatgg

```

<210> 53

<211> 2554

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1796836CB1

<400> 53

```

cggaaatcaa cttccggggg cagaggtggt cgaagccggg tgggtcgtgg gctaccccaa 60
cctgtgtggc tgggcccggg tctccctcca agggcctggg gccgtgcctc ggggtgtacgc 120
gtagggtct gtgtgctgac ggtggctcac cgggcagcgt ggggtgagcgg cgcagcggcg 180
gcagcggaga gcgagagagg ggagcagggt ccacttgaag aatggatgat gatgattttg 240
gtggttttga ggctgcggag acttttgatg gtggaagtgg tgaaccccaa acaacatctc 300
ctgctattcc ttgggctgcc tttcctgcag tatctggagt ccatctttca ccatcttctc 360
ctgagattgt actggaccgt gaccactctt cttccattgg ctgcctctct totgatgcca 420

```

| | | | | | | |
|-------------|------------|------------|------------|-------------|-------------|------|
| ttatttcac | accagagaat | acacatgcag | caaatagcat | tgtgagtc | actattccaa | 480 |
| aagcacagat | tcagcaatca | acacacac | atctggatat | ctcacttttt | ccattgggtt | 540 |
| taactgatga | aaaaagta | ggaacaattg | cccttgtgga | tgattctgag | gatcctggag | 600 |
| ccaatgtatc | taacatacag | cttcagcaaa | aaatttcaag | tctggagatt | aaactcaaa | 660 |
| tatctgaaga | agaaaaacag | agaattaaac | aggatgtgga | atcattgatg | gaaaagcata | 720 |
| atgtcttaga | aaaaggcttt | ctaaaagaaa | aagagcaaga | ggccatttct | tttcaagata | 780 |
| gatacaaa | acttcaggaa | aaacataaac | aagaattgga | agacatgagg | aaagctgggtc | 840 |
| acgaagccct | cagcattatt | gtggatgaat | ataaggcact | actgcagtct | tcagttaagc | 900 |
| aacaagtaga | agctattgaa | aaacagtaga | tttctgcaat | tgagaaacag | gcacacaagt | 960 |
| gtgaggagtt | gctaaatgct | cagcatcaga | ggctccttga | aatgctagat | acagagaagg | 1020 |
| aactgttaaa | agaaaaaata | aaggaagctt | tgattcagca | atctcaagaa | cagaaggaaa | 1080 |
| tattggaaaa | gtgtttggag | gaagaaaggc | aaagaaataa | agaggcatta | gtatccgctg | 1140 |
| caaagcttga | aaaagaagca | atgaaggatg | cagtttttaa | agtcgtagaa | gaagaaagaa | 1200 |
| aaaatttaga | aaaagcgcat | gctgaagaaa | gggaattatg | gaagacagaa | catgcaaaag | 1260 |
| atcaagaaaa | agtatctcag | gaaattcaaa | agctataaca | agaacaaaga | aaaaataagc | 1320 |
| aggaaactgt | taaggcagca | ataatagaag | agcagaaacg | aagtgaagaa | gctgtggaag | 1380 |
| aggcagtga | aagaacaaga | gatgaattga | tagagtatat | aaaagaacag | aaaaggctcg | 1440 |
| atcaagtcac | ccgccaaa | agcctgtcca | gtttggaa | gttctctctc | tgtgcacaga | 1500 |
| aacagttaag | tgctttaata | gctacggaac | cagttgacat | tgataaaaa | gaacatgaca | 1560 |
| aaccacacat | ggcattggat | aaatcatatt | acaccttcaa | aatacacact | ctgaattata | 1620 |
| aagatgtgtt | tgttttcttt | ccaaatcatg | tagaattgat | ttccagttca | aggataaacc | 1680 |
| aaaacaatat | ttagaactat | caagtgtatc | aatttatatt | cttttggttt | cttctttaca | 1740 |
| tttactgtta | ttttattatt | attagtagta | gcagcaacag | agtatgatat | gacccaaaag | 1800 |
| ccattgtaaa | gtgccacatt | accaaaatta | attaagtaaa | ctttatagcc | tgtgggagtc | 1860 |
| tatttatatat | tattttgcaa | aagtagtaaa | tatattattg | tttcatgatg | actcttgatg | 1920 |
| agatgctaga | atgtaaccat | acattttatc | tattttgagg | atagaaatag | catggatttc | 1980 |
| aacatcactt | atttatctgt | ataattggaa | ataaaacacc | gatatgatag | agaatcattc | 2040 |
| cggcattacc | taacctcttc | tcagattgga | tctatgtatt | ttcattgggtc | tactgaaaac | 2100 |
| aaacaataca | attaaaagca | ctaaagatta | ttatattaat | tcaactttga | tctgatata | 2160 |
| cacttaaa | aaaggggtgt | gtgtgggtga | tgcttgtttc | ctatttctgc | tctttaaaga | 2220 |
| tactttgaat | caataaaacc | attagtctac | aaatcaaatt | gtgaacttaa | tctctagaaa | 2280 |
| gagaatataa | ctcagccatt | tataggaatt | taggttcaag | tacaggatat | atgaaatcct | 2340 |
| ttcccagtat | ttcagaatgt | acttaattca | caggcaggat | gcttcaatgc | aaaatcatga | 2400 |
| atatttttaa | ttcaaaacta | aaatgtcatt | aatatgtatg | tatgcaaatg | ttttatctta | 2460 |
| ttttctgaaa | tgcatctact | ttcatgggct | ttgtacgttt | ctgagatttc | tcagtgtaat | 2520 |
| aaaaagagct | cccaaaacta | aaaaaaaaaa | aaaa | | | 2554 |

<210> 54

<211> 1216

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2880670CB1

<400> 54

| | | | | | | |
|------------|-------------|-------------|-------------|-------------|------------|------|
| cgctgttgcc | cttagggacg | gctgtggggc | tgctgggggt | ggggggccga | agcgccagag | 60 |
| atggctgtct | agcgagggat | gcccagctcc | gccgtgaggg | tccctggaaga | ggcggtgggc | 120 |
| atgggtttga | cggcagccgg | ggacgcgagg | gacacggcgg | acgcggtggc | ggctgagggc | 180 |
| gcctactacc | tggaacaggt | caccataact | gaagcatctg | aagatgacta | tgaatatgaa | 240 |
| gagataccag | atgacaattt | tagcatccca | gaaggtgaag | aagatctggc | aaaagcaatt | 300 |
| catagggccc | aagaacaggc | tacagatact | gaaatttttg | aacggaaaac | agttcttctc | 360 |
| tcaaagcatg | cagtacctga | agtaatagaa | gactttctct | gcaatttctt | gatcaaaatg | 420 |
| ggaatgacca | gaactcttga | ttgctttcag | tctgaatggg | atgagttaat | acagaaagga | 480 |
| gtgactgaac | ttagaactgt | tggaatgttt | ccagatgtct | acaccagat | tatgcttttg | 540 |
| gaaaatgaga | acaaaaattt | aaagaaagat | ttgaagcact | acaaacaagc | agctgagtat | 600 |
| gttatttttt | aaatgacatt | ttcttctttt | tcttttggac | taaataaaag | agttgagtga | 660 |
| agctgatata | tgtaatatat | cagagcctta | atttttgaaa | actgaatttt | tctagtgtga | 720 |
| aagaatgtga | gaggcttcat | tagcaaatta | attaaacaga | tgatcagaac | tatcacaatt | 780 |
| ataacttacc | aacaagaagg | gaatgcaggt | agttgttttag | gagatgggtac | atttttgata | 840 |
| taaaattcac | ttccttgtgt | atttgatagt | cttttctatg | tttataacat | tttctcctgt | 900 |
| aaagataggg | taatttctga | aataataatt | aaattttatg | aaagccgaga | ggaaattgct | 960 |
| agtttattcc | tggttagagga | atcttctgtat | ttgaaaattc | tccagaagga | ataatataaa | 1020 |
| ctgtggactt | tggttgataa | tgatatgtag | gttcgtcag | tgtaacaaa | tgtatctctc | 1080 |
| tggtgggggc | tattgataat | ggggaaggct | gtgcagtgtg | gggagtagga | ggtgtatggg | 1140 |
| acatctctgt | accttcta | caattttgct | atgaacttaa | aactgctcta | gaaataaagt | 1200 |

ttattaaaaa aaaaaa

1216

<210> 55

<211> 1457

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2913976CB1

<400> 55

| | | | | | | |
|-------------|------------|-------------|-------------|-------------|-------------|------|
| gaattacagg | tttgagccac | tgcacctggc | caggaacctg | ttttttaacc | ctgggccctt | 60 |
| aacagaggct | gtgtagaaat | ggggggcgtc | acaggccaag | ggcttagtcc | acttcagtcc | 120 |
| agcatccatg | gggaagcctg | gttgccaggag | agtccttcta | agtaggacaa | ggagcagtg | 180 |
| ctgattggag | gacccagtc | agctgccact | gtgaccttcc | tgctgagaga | cttacgcagg | 240 |
| tcccttgctt | ccctgggcct | tggtttctct | ctcggtacaa | tgccgctaata | tgtacctgtt | 300 |
| accacccaga | tgaggaacag | agctgggagg | gctgcatggg | gtcccactgg | atgactcccc | 360 |
| gtctgcatcc | ccacaggccc | cagcggagg | accttcagtc | cagcccggtc | ccctcaggcc | 420 |
| catggaggaa | gagctgccac | ctcccccggc | agaacctgtt | gagaaagggg | catccacaga | 480 |
| catctgtgcc | ttctgccaca | agaccgtgtt | cccccgagag | ctggctgtgg | aggccatgaa | 540 |
| gaggcagtac | catgcccagt | gcttcacgtg | cgcacacctg | cgcgcgccagc | tggctgggca | 600 |
| gagcttctac | cagaaggatg | ggcgaccctt | ctgcgaacct | tgctaccagg | acacactgga | 660 |
| gaggtggcatg | aagtgtggcg | agggtgtccg | ggaccacatc | atcaggggcc | tgggccaggc | 720 |
| cttcaccccc | tcctgcttca | cgtgtgtgac | ctgcgcccgg | tgcattgggg | atgagagctt | 780 |
| tgccctgggc | agccagaacg | agggtgtactg | cctggacgac | ttctacagga | aattcgcccc | 840 |
| cgtctgcagc | atctgtgaaa | atcccatcat | ccctcgggat | gggaaagatg | ccttcaaaaat | 900 |
| cgaatgcata | ggaagaaact | tccatgaaaa | ttgtacagga | tgtaggagct | gcaggatcct | 960 |
| cctgtctgtc | gagcccacgg | accaaggctg | ctacccccctg | aacaaccatc | tcttctgcaa | 1020 |
| gccatgccat | gtgaagcgga | gtgctgcggg | gtgctgtctga | gagtgcctgc | tgggcagtga | 1080 |
| acagaccact | agccccgggt | ggggcccttc | acccgccttg | gcctcccaaa | gtgctggaat | 1140 |
| tacagggtgtg | agccactgtg | cccagctgag | taaatttctt | gattgcacag | aatgtacggt | 1200 |
| gatattggcg | gacttaagga | catcgaaattg | tttatcagga | ataaagtatt | atgtgtgttt | 1260 |
| tctgcggccc | tggataatgc | tgtagcatte | agggtcgatt | gagtaaaaaa | aatgtagaga | 1320 |
| tggtttctctg | tttgtttttg | cttctagtct | tcattcattt | gctattttatt | ctctctctggc | 1380 |
| tttgccttgtg | tatgcatata | tataaaacca | ttattattat | tatcttttagt | ttctagtggg | 1440 |
| aggcttttaa | aaaaaaa | | | | | 1457 |

<210> 56

<211> 1636

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3092084CB1

<400> 56

| | | | | | | |
|------------|------------|------------|-------------|-------------|------------|------|
| ggcgggagct | caggaccggc | gccttctctt | tgttcttggg | ggctcgtggc | ttgctccgc | 60 |
| tgtgcgggaa | aagaatccag | gcccttcac | gcgcgtgtgg | gtgcgggggc | cccgaagtgc | 120 |
| tcgtgggttc | ccgttaggtc | tcgctgggg | caggaaccgg | aatcatgggt | gggaccacca | 180 |
| gcacccgcgc | ggtcaccttc | gaggcggacg | agaatgagaa | catcacctgt | gtgaagggca | 240 |
| tccggctttc | ggaaaatgtg | attgatcgaa | tgaaggaaatc | ctctccatct | ggttcgaagt | 300 |
| ctcagcggta | ttctggtgct | tatggtgcct | cagtttctga | tgaagaattg | aaaagaagag | 360 |
| tagctgagga | gctggcattg | gagcaagcca | agaaagaatc | cgaagatcag | aaacgactaa | 420 |
| agcaagccaa | agagctggac | cgagagagg | ctgctgccaa | tgagcagtta | accagagcca | 480 |
| tccttcggga | gaggatatgt | agcgaggagg | aacgcgctaa | ggcaaagcac | ctggctaggc | 540 |
| agctggaaga | gaaagaccga | gtgctaaaga | agcaggatgc | attctacaaa | gaacagctgg | 600 |
| ctagactgga | ggagaggagc | tcagagttct | acagagtcac | cactgaacaa | tatcagaaag | 660 |
| ctgctgaaga | ggtggaagca | aagttcaagc | gatattgagtc | tcattccagtc | tgtgtgatc | 720 |
| tgaggcccaa | aattcttcag | tgttaccgtg | agaacaccca | ccagaccctc | aaatgctccg | 780 |
| ctctggccac | ccagtatatg | cactgtgtca | atcatgccaa | acagagcatg | cttgagaagg | 840 |
| gaggataaaa | actttcagaa | tgagcaaaac | accatcaacg | ttaattccag | agatggaaag | 900 |
| tttttttttc | tagtgagaaa | acaaccatt | tgaagagaag | accactaatg | agaagaccac | 960 |
| taaagagaga | catcaagaat | ggattcagca | gaatcatttc | acgttttgaa | cagcagcagt | 1020 |
| ttgaagggcc | aaagccttga | tcagggatca | gtcattaaag | gacactcttg | agtattagta | 1080 |
| aaccctctta | tgatgattaa | aagagaagg | cagccctctc | caccttttgg | tactttctat | 1140 |

```

tcaacttgca ctgaccataa aatgtttctc ttctgaacaa gccccatcat ttggtgaacc 1200
tccaccctaa caaagtagga tgggggttgg ggctaaatta attggagtgg ggcgaggaga 1260
gagccagaaa acatagatcc gagggcagca gtgctgggtg gagagagcca gaaaacagat 1320
ctggaggcag cagtgcctga tgggaattgtc taggctgtgg catggttggtt ttgtctttct 1380
tttctccttt gattatgtaa gagctatttc attataactt attatggtga ttatacaggc 1440
aagaagacaa aaaggagaga aaatgtacct cttctactgg aataatgttt atgattacaa 1500
gtgagataag gtatttttat caatatgaag gcaaccttgg ctgataaaac ctctatagtg 1560
aatactcaca tctttacttc actcactatc aataataaat atattttctg acaaagaaaa 1620
aaaaaaaaa aaaaaa 1636

```

```

<210> 57
<211> 1742
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 3882482CB1

```

```

<400> 57
gctgccatgg caaactacg ccgcgcgcgc cttttcacgc acgtgcgag ctaacggact 60
ggcggcgccg ggcgcgccgc gcctgcgcgc caccgcgacc ccatctggac cgcctcgctg 120
aatgtgcccg gacctgcgcc ttctgggtct ctgaaagaag atgaatttgg ctgagatttg 180
tgataatgca agaaaaggaa gagaatatgc ccttcttggg aattacgact catcaatggg 240
atattaccag ggggtgatgc agcagattca gagacattgc cagtcagtca gagatccagc 300
tatcaaaggg aaatggcaac aggttcggca ggaattattg gaggaatatg aacaagttaa 360
aagtattgtc agcactttag aaagttttaa aattgacaag cctccagatt tcctgtgtgc 420
ctgtcaagat gaaccattta gagatcctgc tgtttggcca cccctgttc ctgcagaaca 480
cagagctcca cctcagatca ggcgtcccaa tcgagaagta agacctctga ggaaagaaat 540
ggcaggagta ggagcccgcc gacctgtagg ccgagcacat cctatatcaa agagtgaana 600
gccttctaca agtagggaca aggactatag agcaagaggg agagatgaca aggggaaggaa 660
gaatatgcaa gatggtgcaa gtaatggtga aatgccaaaa ttgatgggtg ctggttatga 720
taaggatctg gtggaagccc ttgaaagaga cattgtatcc aggaatccta gcattcattg 780
ggatgacata gcagatctgg aagaagctaa gaagttgcta agggaagctg ttgttcttcc 840
aatgtggatg cctgactttt tcaaagggat tagaaggcca tggaaggggtg tactgatggg 900
tgaccccccg ggcactggta aaactatgct agctaaagct gttgccactg aatgtgggtac 960
aacattcttc aacgtttcgt cttctacact gacatctaaa tacagagggtg aatctgagaa 1020
gttagtttgt ctgttgtttg agatggctag attttatgcc cctaccacga tcttcattga 1080
tgagatagat tctatctgca gtcgaagagg aacctctgat gaacatgagg caagtccgag 1140
ggtcaagtct gaactgctca ttcagatgga tggagttgga ggagcttttag aaaatgatga 1200
tccttccaaa atggttatgg tattggctgc tactaatctc ccgtgggaca ttgatgaagc 1260
tttgcaaga aggttagaaa aaaggatata tatacctctc ccaacagcaa aagggaagagc 1320
tgagcttctg aagatcaacc ttctgtagggt cgaattagat cctgatattc aactggaaga 1380
tatagccgag aagattgagg gctattctgg tgctgacatc actaatgttt gcagggatgc 1440
ctctttaatg caaatgagac ggcgtatcaa tggcttaagt ccagaagaaa tccgtgcact 1500
ttctaaagag gaacttcaga tgcctgttac caaaggagac tttgaattgg ccctaaagaa 1560
aattgctaag tctgtctctg ctgcagactt ggagaagtat gaaaaatgga tggttgaatt 1620
tggtatctgt tgaatttctg tcagctcttt aatttctggt atttttgttg ataaaaatagc 1680
aagaaattcc tgcaattttt aaaaaacaag tttggaattt tttcagtgga gtggttttgc 1740
ct 1742

```

```

<210> 58
<211> 602
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 4933451CB1

```

```

<400> 58
ctgagttcgg cagctcacac agcaccctt tcaagttggt ttgcagtggg tgcctcttgt 60
gagacactgt ttctgagagc agctttttgt gcatcttaca gggcagattt ctggtccac 120
ccacctctgc ctccgccatg gctgcaagaa ccgttatcat tgaccacggg tctggctttt 180
tgaaaggctg cagggccggc tggaaatgag ctcagatggg cttcccgaac atcgtgaact 240
acctaccgtg caaggagaa cctggcccca gctatgcccg taagcgtgtg agcctgggca 300
tcgacatttg ccatcctgac acctttagct accccatcga gcggggccgc atcctcaact 360

```

```

gggaggggtgt gcagtacctc tgggtcatttg tgttggagaa ccacagacgg gagcaagagg 420
tccccctgt gatcatcacg gagacacctc tgagggagcc tgccgaccga aagaagatgt 480
catccttgga aacctgcag gggacagttt ttccagggtg gcctatcatt ggggtatgag 540
tggctgactg ccctctccat gcctgagagc cactgatttt tcattggcat tccccctggt 600
tt 602

```

<210> 59

<211> 3237

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5043904CB1

<400> 59

```

gaaaatgtct gatgctcagg gcagctacaa actggatgaa gctcaggctg tcttgagaga 60
aacaaaagcc atcaaaaagg ctattacctg tggggaaaag gaaaagcaag atctcattaa 120
gagccttgcc atgttgaagg acggcttctg cactgacagg gggctctcact cagacctgtg 180
gtccagcagc agctctcttg agagtteag tttcccgcta ccgaaacagt acctggatgt 240
gagctcccag acagacatct cgggaagctt cggcatcaac agcaacaatc agttggcaga 300
gaaggtcaga ttgcgccttc gatatgaaga ggctaagaga aggatcgcca acctgaagat 360
ccagctggcc aagcttgaca gtgaggcctg gcctgggggtg ctggactcag agagggaccg 420
gctgatccct atcaacgaga aggaggagct gctgaaggag atgcgcttca tcagcccccg 480
caagtggacc cagggggagg tggagcagct ggagatggcc cggaagcggc tggaaaagga 540
cctgcaggca gcccgggaca ccagagcaa ggcgctgac gagagggtta agttaaacag 600
taagaggaac cagcttgtga gagaactgga ggaagccacc cggcagggtg caactctgca 660
ctcccagctg aaaagtctct caagcagcat gcagtccctg tctcaggca gcagccccgg 720
atccctcacg tccagccggg gctccctggt tgcattccag ctggactcct ccacttcagc 780
cagcttccact gacctctact atgacctctt tgagcagctg gactcagagc tgcagagcaa 840
ggtggagttc ctgctcctgg agggggccac cggcttccgg ccctcaggct gcatcaccac 900
catccacgag gatgaggtgg ccaagaccca gaaggcagag ggaggtggcc gctgcaggc 960
tctgcgttcc ctgtctggca ccccaaagtc catgacctcc ctatccccac gttcctctct 1020
ctcctcccc tccccacct gttccctct catggtgac cccctcctgg ctggtgatgc 1080
cttctcaac tccttggagt ttgaagacc ggagctgagt gccactcttt gtgaactgag 1140
ccttggtaac agcgccagg aaagataccg gctggaggaa ccaggaacgg agggcaagca 1200
gtcgtccaa cctgtgaata cggcccagg gtgtggcctg aaagtggcct gtgtctcagc 1260
cgccgtatcg gacgagtcag tggctggaga cagtgggtgt tacgaggctt ccgtgcagag 1320
actgggtgct tcagaagctg ctgcatttga cagtgcagaa tcggaagcag tgggtgcgac 1380
ccgaattcag attgccctga agtatgatga gaagaataag caatttgcaa tattaatcat 1440
ccagctgagt aacctttctg ctctgttgca gcaacaagac cagaaagtga atatccgct 1500
ggctgtcctt ccttgcctg aaagcacaac ctgcctgtt cggaccggc ctctggagc 1560
ctcagacact ctagtgttca atgaggtgtt ctgggtatcc atgtcctatc cagcccttca 1620
ccagaagacc ttaagagtcg atgtctgtac caccgacagg agccatcttg aagagtgcct 1680
gggagggcgc cagatcagcc tggcggaggt ctgccgtct ggggagaggt cgactcgctg 1740
gtacaacctt ctcagctaca aatacttgaa gaagcagagc agggagctca agccagtggg 1800
agttatggcc cctgcctcag ggcctgccag caaggacgct gtgtctgctc tgttgaaca 1860
gacagcagtg gagctggaga agaggcagga gggcaggagc agcacacaga cactggaaga 1920
cagctggagg tatgaggaga ccagtgaaga tgaggcagta gccaggaag aggaggagga 1980
ggtggaggag gaggaggag aagaggatgt ttccaccgag aaagcctcac ctgatatgga 2040
tgggtaccca gcattaaagg tggacaaaga gaccaacacg gagacccgg ccccatcccc 2100
cacagtgggt cgacctaaag accggagagt gggcaccctg tcccaggggc cattcttcg 2160
agggagcacc atcatccgct ctaagacctt ctcccagga cccagagcc agtacgtgtg 2220
ccggctgaat cggagtgata gtgacagct cactctgtcc aaaaagccac ctttcttcg 2280
aaactccctg gagcgacgca gctccggat gaagcggcct tctcgggtca agtcgttgcg 2340
ctccgagcgt ctgatccgta cctcgctgga cctggagtta gacctgcagg cgacaagaac 2400
ctggcacagc caactgaccc aggagatctc ggtgctgaag gagctcaagg agcagctgga 2460
acaagccaag agccacgggg agaaggagct gccacagtgg ttgcgtgagg acgagcgttt 2520
ccgcctgctg ctgaggatgc tggagaagcg gcagatggac cgagcggagc acaagggtga 2580
gcttcagaca gacaagatga tgagggcagc tgccaaggat gtgcacaggc tccgaggcca 2640
gagctgtaag gaacccccag aagttcagtc ttccaggag aagatggcat ttttaccctg 2700
gcctcggatg aatatcccag ctctctctgc agatgacgtc taatcgccag aaaagtattt 2760
cctttgttcc actgaccagg ctgtgaacat tgactgtggc taaagtattt tatgtgtgt 2820
tatatgaagg tactgagtc caagtctctt agtgccttg ttggtttgaa gatgaaccga 2880
ctttttagtt tgggtcctac tgttgttatt aaaaacagaa caaaaacaaa acacacacac 2940
acacaaaaac agaaaacaaa aaaaccagca ttaaaataat aagattgtat agtttgtata 3000
ttaggagtg tatttttggg aaagaaaatt taaatgaact aaagcagtat tgagttgctg 3060

```

```

ctcttctttaa aatcggttag attttttttg gtttgtacag ctccaccttt tagaggctctt 3120
actgcaataa gaagtaaatgc ctgggtgggac ggtaatccta ataggacgtc ccgcacttgt 3180
cacagtacag ctaatttttc ctagttaaca tattttgtac aatattagga atgcacg 3237

```

<210> 60

<211> 3640

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5202390CB1

<400> 60

```

acctggaata aaaaatccct atcgtgggtg tgttgtgtgg cctgttcctg aaaacattga 60
aatcactgta acacttttta aggatcctca tgcggaagaa tttgaagaca aagagtggac 120
atttgtcata gaaaatgaat ccccttctgg tcgaaggaaa gctcttgcta ctacgagcat 180
caatatgaaa cagtatgcaa gccctatgcc aactcagact gatgtcaagt taaaattcaa 240
gccattatct aaaaaagttg tatctgccgc tcttcagttt tcattatctt gcatttttct 300
gaggggaagga aaagccacag atgaagacat gcaaagtttg gctagtttga tgagtatgaa 360
gcaggctgac attggcaatt tagatgactt cgaagaagat aatgaagatg atgatgagaa 420
cagagtgaac caagaagaaa aggcagctaa aattacagag cttatcaaca aacttaactt 480
tttggatgaa gcagaaaagg acttggccac cgtgaattca aatccatttg atgatcctga 540
tgctgcagaa ttaaatcoat ttggagatcc tgactcagaa gaacctatca ctgaaacagc 600
ttcacctaga aaaacagaag actcttttta taataacagc tataatccct ttaaaggagg 660
gcagactcca cagtatttga acccattcga tgagccagaa gcattttgtga ccataaagga 720
ttctcctccc cagtctacaa aaagaaaaaa tataagacct gtggatatga gcaagtaacc 780
ctatgctgat agttctaaaa ctgaagaaga agaattggat gaatcaaatc ctttttatga 840
acctaaatca actcctcctc caaataatct ggtaaactct gttcaagaac tagaaactga 900
aaggcgagtg aaaagaaaag ccccggtccc accagtcctc tcacctaaaa caggagtatt 960
aatgaaaaac acagtttctg caggaaaaga tctctctact tctcctaagc caagccctat 1020
accaagtctt gttttggggc gaaagccaaa tgctagttag tctttgcttg tatgtgtgaa 1080
agaagttaca aagaactacc gaggagtaaa aatcaccaat tttactacat cgtggagaaa 1140
tggtttatct ttttgtgcaa tattacacca ctttagacca gatttaattg actacaagtc 1200
tctgaatcct caagatatta aagagaacaa caaaaaggca tacgatggat ttgccagcat 1260
aggaatttcc cgattattgg aaccttctga tatggtatta ttagcaattc ctgataaact 1320
gactgttatg acttatctct atcaaaatag ggcacatttc agtggccaag aactaaatgt 1380
cgttcagata gaggaaaaca gcagtaaaag cacatataaa gtgggaaact atgaaacaga 1440
tacaacacag tctgttgatc aagaaaaatt ctatgcagag cttagtgate tgaagcggga 1500
gectgaacta caacagccta tcagcgggag agtagacttc ttatcacagg atgactctgt 1560
atttgaataa ttggagagtg ttggagagtg agaaagttag catcaaactc ctgatgatca 1620
ccttagtcca agcacagcct ccccttactg tcgcaggact aaaagtgaac cagaacccca 1680
gaagtctcag cagagctctg gaaggacttc aggatctgat gacctgggaa tatgttccaa 1740
tacagattca acccaagcac aggttttgtt aggcaaaaag agactattga aagctgagac 1800
tttagaattg agtgacttat atgttagtga taagaagaag gatattgtct caccctttat 1860
ttgtgaggag acagatgaac aaaagcttca aactctagac atcggtagta acttggagaa 1920
agaaaaatta gagaattcca gatccttaga atgcagatca gatccagaat ctccatcaa 1980
aaaaacaagt ttatctccta cttctaaact tggatactca tatagtagag atctagacct 2040
tgctaagaaa aaacatgctt ccttgaggca gatccagatg gatccagatg ctgatagaac 2100
cactttaaat catgcagatc attcatcaaa aatgattccag catcgattgt tatctagaca 2160
agaagaactt aaggaaagag caagagtctt gcttgagcaa gcaagaagag atgcagcctt 2220
aaaggcgggg aataagcaca ataccaacac agccacccca ttctgcaaca ggcagctaa 2280
tgatcagcaa gatgaagagc gacgtcggca gctgagagag agagctcgtc agctaatag 2340
agaagctcga tctggagtga agatgtcaga acttcccagc tatggtgaaa tggctgcaga 2400
aaagtgtgaa gaaaggtcaa aggcattctg agatgaaaat gataatattg agatagatac 2460
taacgaggag atccctgaag gctttgttgt aggaggtgga gatgaactta ctaacttaga 2520
aatgacctt gatactcccg aacaaaacag taagttgggt gacttgaagc tgaagaagct 2580
cctagaagtt cagccacagg tggcaaattc accctccagt gctgcccaga aagctgtaac 2640
tgagagctca gagcaggaca tgaaaagtgg cacagaagat ctccggactg aacgattaca 2700
aaaaacaaca gaacgtttta gaaatcctgt tgbgttcagc aaagattcta cagtcagaaa 2760
aactcaactt cagtctttca gccaatatat tgagaataga ccagagatga aaaggcagag 2820
atcaatacag gaagatacaa agaaaggaaa tgaggagaag gcagcgataa ctgaaactca 2880
gaggaaagcca tcagaagatg aagtgtctaa taaaggggtc aaagacacca gtcagtatgt 2940
agtaggagaa ttggcagcac tagagaatga gcaaaagcaa attgacaccc gtgccgcgct 3000
ggtggagaag cgccttcgct atctcatgga cacaggaagg aacacagaag aagaagaagc 3060
tatgatgcag gaatggttta tgttagttaa taagaaaaat gccttaataa ggagaatgaa 3120
tcagctctct cttctggaag aagaacatga tttagaacga cggtatgagc tgctgaaccg 3180

```

| | | | | | | |
|------------|------------|------------|-------------|------------|------------|------|
| ggaattgagg | gcaatgctag | ccattgaaga | ctggcagaag | accgaggccc | agaagcgacg | 3240 |
| cgaacagctt | ctgctagatg | agctgggtgc | cctgggtgaac | aagcgcgatg | cgctcgtcag | 3300 |
| ggacctggac | gcgcaggaga | agcaggccga | agaagaagat | gagcatattg | agcgaactct | 3360 |
| ggagcaaaac | aaaggcaaga | tggccaagaa | agaggagaaa | tgtgttcttc | agtagccatc | 3420 |
| agatcagaaa | gaatctctcc | caacatttta | gagtcttgc | tcccaaacca | gaaaaagtca | 3480 |
| gactcattgt | tgatttaaaa | cttttaacat | tttgtttggc | tggattgtac | tactttacct | 3540 |
| ctactttacc | accaccaccc | ttttcctccc | tcctttccaa | ataatatata | gaactccaaa | 3600 |
| atagcttcat | ttaaggattt | ttttgtgagt | taacaatttc | | | 3640 |

<210> 61
 <211> 2111
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5526375CB1

| | | | | | | |
|-------------|-------------|-------------|------------|-------------|-------------|------|
| <400> 61 | | | | | | |
| gcaggacatg | gaagaacgcg | gggtcacccga | cggggatctc | gcgcggagcc | tggagcaagg | 60 |
| gccagagggg | ccggaacgc | ccatccaggt | ggtgctcagg | gtacgtccca | tgagcgcggc | 120 |
| cgagctgctg | cgagggcagc | agagcgtgct | gcactgctca | gggaccggga | ctctgcagggt | 180 |
| gagtcctcca | ggcgggggtc | cagaagtggc | gttccgcttc | ggtgcgggtgc | tagacgcggc | 240 |
| gcgcacgcag | gaggacgtgt | tccgggcgtg | cggcgtgcgg | cgccctggggg | agctggcgct | 300 |
| gcgcgggtttc | tcctgcactg | ttttcacctt | tggccagacg | ggctctggga | agacctacac | 360 |
| cctgactgga | ccccctcccc | agggggaggg | ggtgcctgta | ccccccagcc | tggctggcat | 420 |
| catgcagagg | accttcgcct | ggctgttgga | ccgcgtgcag | cacctgggtg | ccctgtcac | 480 |
| ccttcgcgccc | tcttatctgg | agatctacaa | tgggcagggt | cgggacttgc | tgagcctggg | 540 |
| gtctccccgg | ccccctccctg | ttcgctggaa | caagactcgg | ggcttctatg | tggagcagct | 600 |
| gcgggtgggtg | gaatttggga | gtctggaggc | cctgatggaa | cttttgcaaa | cgggtctcag | 660 |
| ccgtcgaagg | aactcagccc | acaccctgaa | ccaggcctcc | agccgaagcc | atgccctgct | 720 |
| caccctttac | atcagccgtc | aaactgccc | gcagatgcct | tctgtggacc | ctggggagcc | 780 |
| ccctgttggt | gggaagctgt | gctttgtgga | cctggcaggc | agtgagaagg | tagcagccac | 840 |
| gggatcccg | ggggagctga | tgcttgaggc | taacagcatc | aaccgaagcc | tgctggccct | 900 |
| gggtcactgc | atctccctgc | tgctggaccc | acagcggaa | cagagccaca | tcctttccg | 960 |
| ggacagcaag | ctcaccaagt | tgctggcaga | ctcactggga | gggcgcgggg | tcaccctcat | 1020 |
| ggtggcctgc | gtgtccccct | cagcccagtg | ccttcctgag | actctcagca | ccctgcgata | 1080 |
| tgcaagccga | gctcagcggg | tcaccaccgg | accacaggcc | cccaagtctc | ctgtggcaaa | 1140 |
| gcagccccag | cgtttgagga | cagagatgct | gcagctccag | gaggagaacc | gtcgcctgca | 1200 |
| gttcacagctg | gaccaaattg | actgcaaggc | ctcagggtc | agtggagccc | gggtggcctg | 1260 |
| ggcccagcgg | aacctgtacg | ggatgtctaca | ggagttcatg | ctagagaatg | agaggctcag | 1320 |
| gaaagaaaag | agccagctgc | agaatagccg | agagctagcc | cagaatgagc | agcgcatcct | 1380 |
| ggcccagcag | gtccatgcac | tagagaggcg | tctcctctct | gcctgtctacc | atcaccagca | 1440 |
| gggtcctggc | ctgacccac | cgtgtccctg | cttgatggcc | ccagctcccc | cttgccatgc | 1500 |
| actgccaccc | ctctactcct | gcccctgtcg | ccacatctgc | ccactgtgtc | gagtgccct | 1560 |
| ggcccactgg | gctgcctgc | caggggagca | ccacctgccc | cagcctctct | tctgggtctc | 1620 |
| gaggagtcag | aaatagacca | gacgtgggtc | ctgctctcag | gaggcttcta | gtctcaggag | 1680 |
| aggacggtag | aagaaccatc | tgctgcacac | agaaggcgac | caagacccca | cccctttatg | 1740 |
| ccatcctgac | attcttggtg | ccttgtgggg | taggcatggt | ggggtgtcat | ggctagcacc | 1800 |
| aggtaggtgg | gacacagaga | ggcaggactg | aggtcaccca | gagtggactc | tggtctttcc | 1860 |
| tacaggtgtt | ggaccctgag | gcctcagggtg | gcaggccccc | atctgcccgg | ccccaccct | 1920 |
| gggcaccccc | atgcagccct | ggctctgcca | agtgcccaag | agagaggagt | cacagtgcac | 1980 |
| ggactcagac | ccgagtcctg | gcagagatgt | tgacggagga | ggaggtggta | ccttctgcac | 2040 |
| ctccccctgc | tgtgagggcc | ccgaagacat | caccagggtc | cagaggtggg | gcccgggttc | 2100 |
| aaacctggcc | c | | | | | 2111 |

<210> 62
 <211> 1389
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5677408CB1

| | | | | | |
|------------|------------|-----------|-----------|------------|------------|
| <400> 62 | | | | | |
| cccagactcc | ttgcggactc | gcccctgat | tctaggtgg | tcactactcc | gagcctgtga |
| | | | | | 60 |


```

cgtttgccgc agccaggccg tcgacgatgc ccagtgaaac tctctgggaa attgcaaaaag 120
ctgaagtggg aaaaaggggg attaatggaa gtgaagggtg tggagctgaa attgcagaaa 180
aatttgtttt cttcattggc agtaaaaatg ggggaaagac tactattatt ctaagggtgtc 240
ttgacagaga tgaaccacca aaaccaacct tagcttttga atatacatat ggaagaagag 300
caaaagggca caacacacca aaagatatcg ctcacttttg ggaactcggg ggaggaacct 360
ctttatttga cttaatcagc ataccatca caggtgacac cttacggacg tttctcttg 420
ttctcgttct ggatctttca aaacctaatg atctctggcc caccatggaa aatctcttgc 480
aagccacaaa aagccatgta gacaaagtga taatgaaact gggaaagaca aatgctaaag 540
cagtttctga aatgagacag aagatcttga ataatatgcc gaaggatcat cctgatcatg 600
aattaattga cccatttccg gtacctctgg tcataattgg aagtaaatat gatgtttttc 660
aggattttga gtctgagaag agaaaagtaa tatgcaagac acttcgattt gttgcacatt 720
attatggagc atcattaatg tttaccagta aatcagaagc tctattacta aaaatacgtg 780
gagttatcaa ccagttggca tttggcattg acaaaagcaa atcaatatgt gtggatcaga 840
ataaaccgct gtttatcaca gcaggattgg attctttcgg tcaaatagga tctcctcctg 900
ttcctgaaaa tgacatttga aagcttcatg cccactcacc tatggagtgg tggaaaaaag 960
tgtatgaaaa ctcctttcca ccaaagagta ttaacacgct gaaagatatac aaggaccctg 1020
cgagagatcc tcagtatgct gaaaatgaag tcgatgagat gagaattcag aaggatctgg 1080
aactggaaca gtacaaaaga agttcttcca agtcttggaa acaaatcgag cttgattctt 1140
gaacctattt caattattgt atattttatt cttcttttcc aaatacaaat aagattatac 1200
tgtgaattaa ctattgtggc aatatgtgaa gaaagttaaa ctgtataatt tgtaaaggga 1260
caagctggat ttcttggaat agtgcacatc cctgtataat ttgaagcttt ttaaaaggaa 1320
aaattattgt agaaccacgt gtaatttttt ttaaaataaa agaattctct actacctaca 1380
aaaaaaaaa

```

<210> 63

<211> 3331

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5982278CB1

<400> 63

```

gtcgccgcgc gcttcgcaga gcaccgcgcc ttagecgcga agttctagtt cttgctgccg 60
gtcctaacgt cccgagctct tcgccagcca gccgtcccgc atgcgcgttt gggcgccgtg 120
gagcctgctg ccatgaagtc agcgagagct aagacacccc ggaaacctac cgtgaaaaaa 180
gggtcccaaa cgaaccttaa agaccaggt ggggtatact gtagggtgag cccactgggc 240
tttctgatac aagagtgttg catagaagtg atcaataata caactgttca gcttcatact 300
cctgaggggt acagactcaa ccgaaatgga gactataagg agactcagta ttcattttaa 360
caagtatttg gcactcacac caccagaag gaactctttg atgttgtggc taatcccttg 420
gtcaatgacc tcattcatgg caaaaatggt cttcttttta catatggtgt gacgggaagt 480
ggaaaaactc acacaatgac tggttctcca ggggaaggag ggctgcttcc tcgttgtttg 540
gacatgatct ttaacagtat agggtcattt caagctaaac gatatgtttt caaatctaat 600
gataggaata gtatggatat acagtgtgag gttgatgcct tattagaacg tcagaaaaaga 660
gaagctatgc ccaatccaaa gacttcttct agcaaacgac aagtagatcc agagtttgca 720
gatatgataa ctgtacaaga attctgcaaa gcagaagagg ttgatgaaga tagtgtctat 780
ggtgtatttg tctcttatat tgaaatatat aataattaca tatatgatct attggaagag 840
gtgccgtttg atcccataaa cccaaacctc cacaatctaa attgcttcgt gaagattaag 900
aaccataaca tgtatgttgc aggtgtgac gaagttgaag tgaaatctac tgaggaggct 960
tttgaagttt tctggagagg ccagaaaaag agacgtattg ctaataccca tttgaatcgt 1020
gagtcagccc gttcccatag cgtgttcaac attaaattag ttcaggctcc cttggatgca 1080
gatggagaca atgtcttaca ggaaaaagaa caaatcacta taagtcagtt gtccttggta 1140
gatcttgctg gaagtgaag aactaaccgg accagagcag aagggaacag attacgtgaa 1200
gctggtaata ttaatcagtc actaatgacg ctaagaacat gtatggatgt cctaagagag 1260
aaccaaatgt atggaactaa caagatggtt ccataatcag attcaaagtt aacctatctg 1320
ttcaagaact actttgatgg ggaaggaaaa gtgcggatga tcgtgtgtgt gaaccccaag 1380
gctgaagatt atgaagaaaa cttgcaagtc atgagatttg cggaagtgac tcaagaagtt 1440
gaagtagcaa gacctgtaga caaggcaata tgtggtttta cgctgggag gagatacaga 1500
aaccagcttc gagggtccag tggaaatgaa ccattgggta ctgacgtggg tttgcagagt 1560
tttccacctt tgccgtcatg cgaaattttg gatatacaac atgagcagac acttccaagg 1620
ctgattgaag ccttagagaa acgacataac ttacgacaaa tgatgattga tgagtttaac 1680
aaacaatcta atgcttttaa agcttttgta caagaatttg acaatgctgt ttaagttaa 1740
gaaaaccaca tgcaagggaa actaaatgaa aaggagaaga tgatctcagg acagaaattg 1800
gaaatagaac gactggaaaa gaaaaacaaa actttagaat ataagattga gattttagag 1860
aaaacaacta ctatctatga ggaagataaa cgcaatttgc aacaggaact tgaaactcag 1920
aaccagaaac ttcagcgaca gttttctgac aaacgcagat tagaagccag gttgcaaggc 1980

```

| | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| atggtgacag | aaacgacaat | gaagtgggag | aaagaatgtg | agcgtagagt | ggcagccaaa | 2040 |
| cagctggaga | tgagaataaa | actctgggtt | aaagatgaaa | agctgaaaca | actgaaggct | 2100 |
| attgttactg | aacctaaaac | tgagaagcca | gagagaccct | ctcgggagcg | agatcgagaa | 2160 |
| aaagttactc | aaagatctgt | ttctccatca | cctgtgcctt | tactctttca | acctgatcag | 2220 |
| aacgcaccac | caattcgtct | ccgacacaga | cgatcacgct | ctgcaggaga | cagatgggta | 2280 |
| gatcataagc | ccgcctctaa | catgcaaact | gaaacagtca | tcgagccaca | tgtccctcat | 2340 |
| gccatcacag | tatctgttgc | aaatgaaaag | gcactagcta | agtgtgagaa | gtacatgctg | 2400 |
| accaccagg | aactagcctc | cgatggggag | attgaaacta | aactaattaa | gggtgatatt | 2460 |
| tataaaacaa | gggttggttg | acaatctgtt | cagtttactg | atattgagac | tttaaagcaa | 2520 |
| gaatcaccaa | atggtagtgc | aaaacgaaga | tcttccacag | tagcacctgc | ccaaccagat | 2580 |
| ggtgcagagt | ctgaatggac | cgatgtagaa | acaagggtgt | ctgtggctgt | ggagatgaga | 2640 |
| gcaggatccc | agctgggacc | tggatatcag | catcacgcac | aaccaagcg | caaaaagcca | 2700 |
| tgaactgaca | gtccagctac | tgaagaaca | ttttcatttg | tgtggatgat | ttctcgaaag | 2760 |
| ccatgccaga | agcagtcctc | caggtcatct | tgtagaactc | cagctttgtt | gaaaatcacg | 2820 |
| gacctcagct | acatcataca | ctgaccacga | gcaaagcttt | ccctatgggt | ccaaagacaa | 2880 |
| ctagtattca | acaaaccttg | tatagtgtat | gttttgccat | atttaattat | aatagcagag | 2940 |
| gaagactcct | tttttcatca | ctgtatgaat | tttttataat | gtttttttta | aatatatttc | 3000 |
| atgtatactt | ataaactaat | tcacacaagt | gtttgtctta | gatgattaag | gaagactata | 3060 |
| ctagatgact | gtctgatttt | ttattgtgac | ttctccagcc | ctggtctgaa | ttctctagg | 3120 |
| ttttataaac | aatgtctgct | atattatgac | tgcaagaatg | cacttttaga | ctatttgaca | 3180 |
| attcagactt | tcaaaataaa | gatgtaaatg | actggccaat | aataaccatt | ttaggaaggt | 3240 |
| gttttgaatt | ctgtatgtat | atattcactt | tctgacattt | agatatgcc | aaagaattaa | 3300 |
| aatcaaaagc | actaagaaat | acaaaaaaaa | a | | | 3331 |

<210> 64

<211> 3558

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 6437362CB1

<400> 64

| | | | | | | |
|-------------|-------------|-------------|------------|-------------|-------------|------|
| ggcggccgctc | ctcatcccgg | cgcttgagag | gacgcggggc | tgcgcaaattg | gcttgtcccg | 60 |
| ctctcgggtct | ggaagctctt | cagccccctgc | agcccccgag | gccccccgag | cccgccttct | 120 |
| ccgaggcgca | gaagtggatt | gagcaagtaa | ctggcagaag | ttttgggtgat | aaagattttc | 180 |
| ggacagggttt | agaaaaatgga | atcctcctct | gcgagttgct | gaatgctata | aagccaggac | 240 |
| ttgttaaaaa | gatcaataga | ttgcctaccc | ccattgcagg | actggacaat | attatcttat | 300 |
| tcttgagagg | ttgtaaagag | ctcggcctta | agaatctca | actttttgac | ccgagtgacc | 360 |
| tccaggatac | atccaacaga | gtaacagtca | agagccttga | ttatagtagg | aagctgaaaa | 420 |
| atgtattagt | taccattttac | tggctgggaa | aagcagcaaa | cagctgcaca | tcctacagcg | 480 |
| gaacgacact | aaacctgaag | gagtttgaag | gattgttggc | tcagatgcga | aaggacactg | 540 |
| atgacattga | aagtcttaaa | cgcagtatcc | gagacagtgg | ctacatcgac | tgttgggatt | 600 |
| ccgagcgcag | cgactccctc | tctcctcctc | gccacggcag | agatgattcc | ttcgacagcc | 660 |
| tggattcctt | tggctctcgc | tctcggcaga | cgccctcacc | agatgtagtc | ctcaggggaa | 720 |
| gcagcgatgg | gagaggaagc | gactctgaat | ccgacttgcc | tcacggaag | ctgcagatg | 780 |
| tgaagaagga | tgacatgtct | gcacggcgga | cttcccattg | tgagccgaaa | tcagcagtcg | 840 |
| cttttaacca | gtacctcccg | aacaaaagca | atcacagcgc | ctacgtcccc | gcgcctctga | 900 |
| gaaagaagaa | agcagagaga | gaggaatacc | gcaagagctg | gagtaaccgc | acctccccgc | 960 |
| tgggtgggga | gaggcccttc | agatacggtc | cgagaactcc | tgtgtctgat | gacgcagaga | 1020 |
| gcaccagcat | gtttgacatg | cggtgtgagg | aggaggccgc | ggtgcagccg | cacagcaggg | 1080 |
| cccgcaggga | gcagctgcag | ctgataaata | accagctgag | ggaagaggac | gacaaatggc | 1140 |
| aagatgacct | ggctcgttgg | aagagtcgta | gaagaagtgt | ttctcaggac | ttaatcaaga | 1200 |
| aagaggaaga | aaggaaaaaa | atggagaagt | tactggctgg | agaagatggg | acaagtgaac | 1260 |
| gaaggaaaag | catcaaaacc | tacagagaaa | ttgttcaaga | aaaagagcgg | agagagagag | 1320 |
| agctgcatga | agcatataag | aacgctcgg | cccaggagga | ggcagagggg | atccttcaac | 1380 |
| agtacattga | gaggttcacc | atcagtgagg | ctgttctcga | acgcttggag | atgccaaaaa | 1440 |
| ttctggaaag | aagccattca | acagagccaa | atttatcctc | cttctgaaat | gaccccaatc | 1500 |
| ccatgaaata | cctgcggcaa | cagtcactgc | ctccacccaa | attcactgcc | actgttgaaa | 1560 |
| ccaccattgc | tcgtgccagt | gttctggata | ccagcatgtc | agcaggcagt | gggtctccaa | 1620 |
| gcaaaactgt | cactcccaaa | gcagtgccta | tgtctgacac | caagccttac | tcccagccca | 1680 |
| aaaattctca | agatgttctg | aagaccttta | aggtagacgg | gaaagtccag | gtgaaatggg | 1740 |
| agacggttca | tagagaggag | gagaaggaaa | gagagtgtcc | cacggtggca | cctgcccact | 1800 |
| ccctaaccac | atcccagatg | tttgaagggt | tggccagagt | gcacgggtct | ccactggagc | 1860 |
| tgaacaaga | caacggtagc | atcgagatca | acataaagaa | gccaaactct | gttcccccaag | 1920 |
| agctcgcagc | aaccactgag | aaaacggaac | cgaatagtca | agaggacaag | aatgatgggtg | 1980 |

| | | | | | | |
|-------------|-------------|-------------|------------|-------------|-------------|------|
| gaaaatcaag | aaaaggggaat | atagaacttg | cctcatcaga | accacagcat | tttacaacaa | 2040 |
| ctgtgactcg | atgcagcccg | accgtggcct | ttgtggaatt | tccctccagc | ccccagctga | 2100 |
| agaatgatgt | gtcggaaaga | aaagaccaga | agaaaccaga | aatgaaatg | agtggaaagg | 2160 |
| tggagttggg | gctgtcacia | aaggtggtaa | agccaaaatc | tccagaaccc | gaagcaacgc | 2220 |
| tgacatttcc | atttctggac | aaaatgcctg | aagccaacca | actacatttg | ccaaatctca | 2280 |
| attctcaagt | ggattctcca | agcagtgaga | agtcacctgt | tacgacacct | tttaagttct | 2340 |
| gggcatggga | cccagaagag | gagcgagcgc | gacaggaaaa | atggcaacag | gaacaggaaac | 2400 |
| gtttgtccca | ggagagatac | cagaaggagc | aggacaagct | gaaagaagag | tgggaaaagg | 2460 |
| ccaaaaggga | ggtggaagag | gaagaacgca | gatactatga | ggaggagcgt | aagataattg | 2520 |
| aagacactgt | ggttccattt | actgtttctt | caagttccgc | tgaccagctg | tctacctctt | 2580 |
| cctccatgac | tgaaggcagt | gggacaatga | ataagataga | cctgggaaac | tgtaaatg | 2640 |
| aaaaacaaga | cagaagatgg | aagaaatcat | tccagggaga | tgacagtgc | ttattgtctga | 2700 |
| agactaggga | aagtgatcga | ctggaggaga | agggcagcct | aactgaaggg | gccttggctc | 2760 |
| attctgggaa | ccctgtatca | aaaggagtcc | atgaagacca | tcagctggat | accgaggctg | 2820 |
| gggccccaca | ctgtggaaca | aaccacacgc | ttgctcagga | tccatcccag | aatcagcaga | 2880 |
| catcaaatcc | aacgcacagt | tcagaagatg | tgaagccaaa | aacctcccg | ctggataaaa | 2940 |
| gcattaacca | tcagatcgag | tctcccagtg | aaaggcggaa | gtctataagt | ggaaagaagc | 3000 |
| tgtgtctctt | ctgtgggctt | cctttgggta | aaggagctgc | aatgatcatc | gagaccctca | 3060 |
| atctctattt | tcacatccag | tggttcaggt | gtggaatttg | taaaggccag | cttggagatg | 3120 |
| cagtgtgtgg | gacggtggtt | aggattcgaa | atggtctcct | gaactgtaat | gattgtctaca | 3180 |
| tgcatccag | aagtgcgggg | cagcctacaa | cattgtgaca | cggctttcaa | gcttccggat | 3240 |
| cactcaccat | ttctttactg | agagtgtccc | ctggcaactg | cttaacaaaa | tccaagctc | 3300 |
| aggggcttct | cagcattttac | ctaatttctg | aaaggctctt | ctgaaagggtg | gtatctgttc | 3360 |
| ttctgtagta | cagtgtttat | gtttttctctg | tttattgttt | tggttttttt | tttttttttg | 3420 |
| cattttgcaca | gtatacacia | aagaatatgg | gggttgaatg | atcctgaata | gctcaaaaaa | 3480 |
| ggtttttagca | tggtcaaaaa | ggcttatggt | ttaaaatgtg | ttattctctt | ctttgggaat | 3540 |
| tagctaaatg | atgcaata | | | | | 3558 |

<210> 65

<211> 5373

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4173970CB1

<400> 65

| | | | | | | |
|-------------|-------------|-------------|------------|-------------|-------------|------|
| cccgactgtc | taatcattag | tctcaaaaat | tattgcttgg | tgaaagcttt | ttttaaaaaa | 60 |
| aagcttcagg | aaacaaaagt | gcatttgggtg | gtatatgaag | ataacctaaa | acaggatata | 120 |
| tgtttcattt | atgggccttt | gtcttattct | ttggtctgag | atgatttttt | tctgagttga | 180 |
| cttttttaggt | gccacgtgtg | ccaacgcata | caattgccat | ggttgtacct | ccccaggaac | 240 |
| ctgacagaac | ttcacaggag | aactctctctg | cccttttagg | agtgcacaaa | gctgtgagta | 300 |
| ccagagtgcc | cactgtgtcc | aacagttctt | ctcagaccac | agagtgtctt | acacctgaat | 360 |
| cctgttcgca | gactacaagc | aatgtggcct | cccaatcgat | gcctctctgtg | tatctctcag | 420 |
| ttgacattga | tgacataact | gagagcaatc | atgacacagc | attaacacta | gcttgtgcag | 480 |
| gtggatcatga | agaacttgta | tctgtgtctca | ttgcacggga | tgccaaaatt | gaacacagag | 540 |
| acaaaaaagg | ttcacacca | ctaactcctgg | cagcaacagc | agggcatggt | ggagttgttg | 600 |
| aaatcctttt | ggataaagg | ggagatatag | aagcacagtc | tgaaacgaact | aaggatactc | 660 |
| cgctttcatt | ggcatgttct | ggtggacgtc | aggaggtggt | agacttgctg | ctggctcgag | 720 |
| gtgcaataaa | agaacatagg | aacgtatctg | attatacacc | actgagtcta | gctgctgtctg | 780 |
| gaggatatgt | taatatacatt | aagattctgc | ttaatgctgg | ggcagaaatt | aattcaaggga | 840 |
| ctgggagtaa | actaggtatt | tctcccctga | tggttgctgc | aatgaatgga | catgttctctg | 900 |
| cagtaaaatt | gctgctcgat | atgggttcag | acattaatgc | ccaaatagag | accaatcgga | 960 |
| acacggctct | caccctggcc | tggttccagg | gccgagcaga | agtagtgagt | ttgcttctgg | 1020 |
| accgaaaagc | caatgttgaa | catagggcaa | agacgggtct | taccccttg | atggaagcag | 1080 |
| cttctggagg | gtatgcagag | gttggaagag | ttcttcttga | taaaggagca | gatgttaattg | 1140 |
| ctccccctgt | gccttctca | agagatactg | ctttaacaat | agcagcagac | aaaggtcact | 1200 |
| acaaaatttg | tgaactcctg | attcatatgg | gagccacat | tgatgttctg | aacaaaaagg | 1260 |
| gaaatacgcc | actttggctg | gcatccaatg | gaggtcattt | tgatgttctg | cagttgctag | 1320 |
| tgcaagcagg | tgctgatgtg | gatgcagcag | ataaccggaa | aatcacacct | cttatgtcag | 1380 |
| catttcgcaa | gggtcatgta | aaagttgttc | aatatttgg | aaaggaagta | aatcagttcc | 1440 |
| cttctgatata | agaatgcag | agatacatag | caacaattac | agataaggaa | ctgttgaaaa | 1500 |
| aatgtcatca | atgtgtcga | accattgtga | aggctaaaga | ccagcaagct | gcagaagcaa | 1560 |
| ataagaatgc | gagtattctt | ttaaaggaa | ttgatctgga | aaagtcaaga | gaagagagca | 1620 |
| gaaagcaggc | tcttgcgtg | aaaagagaaa | aaagaaaaa | aaagagaaaa | aagaaaaaag | 1680 |
| aggaacagaa | aaggaaacag | gaagaagatg | aagaaaaaa | acctaaggag | aattcgggaa | 1740 |

| | | | | | | |
|-------------|------------|-------------|------------|-------------|-------------|------|
| taccagagga | tgaagatgaa | gaggagaaatg | atgaagatgt | ggagcaagaa | gttcccatag | 1800 |
| aacctcctag | tgcaaccacc | accactacga | ttggaatctc | tgcaacatct | gcaacattca | 1860 |
| caaatgtgtt | tggaaaaaaa | agggccaatg | tggtgacaac | tcccagcacc | aatcggaaaa | 1920 |
| ataagaagaa | caaaacaaaa | gaaacccctc | ctacagcaca | tttaatttta | ccagaacaa | 1980 |
| atatgtcttt | agcccaacaa | aaggcagata | aaaataaaat | aaatggagaa | cctagagggtg | 2040 |
| gtggtgcagg | tggaatagat | gattcagata | acttggacag | cacagactgc | aacagtgaga | 2100 |
| gtagcagtg | tggtaaaagc | caagagttaa | atattgtgat | ggatgtgaat | tcctctaaat | 2160 |
| accctcact | gctccttcat | tcccaagaag | aaaagacaag | tactgctact | tccaaaactc | 2220 |
| agacacgcta | caagacagtg | tcattggccat | taagctctcc | aaacataaag | ctgaatctca | 2280 |
| ctagccctaa | aaggggtcag | aaaagagaag | aaggggtgaa | agaagtgtga | cgaagggtcaa | 2340 |
| agaaattgtc | tggtccagcc | tcagtgggtg | cgaggataat | gggaagagga | ggatgtcaaca | 2400 |
| tactgcaat | acaggatgtt | actggtgccc | atattgatgt | ggataaacia | aaagataaga | 2460 |
| atggcgagag | aatgatcaca | ataaggggtg | gcacagaatc | aacaagatat | gcagttcaac | 2520 |
| taatcaatgc | actcattcaa | gatcctgcta | aggaactgga | agacttgatt | cctaaaaatc | 2580 |
| atatcagaac | acctgccagc | accaaataca | ttcatgctaa | cttctcatct | ggagtgtgtg | 2640 |
| ccacagcagc | ttccagttaa | aatgcatttc | ctttgggtgc | tccaactctt | gtaacttcac | 2700 |
| aggaacaaac | gttatctacg | ttccagcccg | ctaataaact | taataagaat | gttccaacaa | 2760 |
| atgtacgttc | ttctttccca | gtttctctac | ccttagctta | tctcaccctt | cattttgccc | 2820 |
| tgctggctgc | tcaaactatg | caacagattc | ggcactcctc | cttaccatgc | gcccagtttg | 2880 |
| gaggaaccta | ctcaccttct | cctaaccatc | ggggaccatt | ccagtgaga | cctgtgaatc | 2940 |
| ctggcaacac | aaatagctct | ccaaagcata | ataacacaag | ccgtctacct | aaccagaacg | 3000 |
| ggactgtttt | acctcagag | tctgctggac | tagctactgc | cagttgtcct | atcactgtct | 3060 |
| cttctgtagt | tgctgccagt | cagcaactgt | gtgtcactaa | taccgggact | ccttcactag | 3120 |
| tcagaaagca | gttgttgcct | tggtgtccta | agacaagtcc | tccagcaaca | gtgatttctt | 3180 |
| ctgtgacaag | cacttgtagt | tccttgccct | ctgtctcttc | tgcacctatc | actagcgggc | 3240 |
| aagctcccac | cacatttcta | cctgcaagta | cttctcaagc | acagctttct | tcacaaaaga | 3300 |
| tggagtcttt | ctctgctgtg | ccaccaccca | aagagaaagt | gtccacacag | gaccagccca | 3360 |
| tggcaaacct | atgtaccca | tcttcaactg | caaacagttg | cagtagctct | gccagcaaca | 3420 |
| ccccgggagc | tccagaaact | caccatccca | gtagtccac | tctacttcc | agtaacacac | 3480 |
| aagaggaggc | acagccatcc | agtgtgtctg | atttaagtcc | tatgtcaatg | ccttttgcat | 3540 |
| ctaactcaga | acctgtccca | ttgactttga | catcaccag | aatggttgct | gctgataatc | 3600 |
| aggacaccag | taatttacct | cagttagctg | taccagcacc | tcgagtttct | catcgatgc | 3660 |
| agccagagg | tcttttttac | tcctgtgtac | caaagcaac | tattcaccag | gatccccagt | 3720 |
| ctatttttgt | tacgaatcca | gttactttta | caccacctca | aggccacca | gctgcagtgc | 3780 |
| agctttcttc | agctgtgaac | attatgaatg | gttctcagat | gcacataaac | ccagcaataa | 3840 |
| agtctttgcc | acctacattt | ggcccagcca | cacttttcaa | tcacttcagc | agtctttttg | 3900 |
| atagtagtca | ggtgccagct | aaccagggtc | ggggagatgg | tccactgtcc | tcacagattg | 3960 |
| ctacagatgc | ctcttttact | gttcagtcag | cgttcctggg | taactcagtg | cttggacact | 4020 |
| tggaaaacat | gcaccctgat | aactcaaagg | cacctggctt | cagaccacct | tcccagcgag | 4080 |
| tttctactag | tccagttggg | ttaccatcca | ttgaccatc | aggcagctcc | ccatcttctc | 4140 |
| cctctgtctc | tctggcaagt | ttttccggga | taccaggaac | aagggttttc | ctgcaagggc | 4200 |
| cagctcctgt | tgggactcct | agtttcaaca | gacaacattt | ttctcccat | ccttggacaa | 4260 |
| gcgcctcaaa | ctcatgtgac | tctcctatct | catctgtttc | ttcgggatca | tcttcacctc | 4320 |
| tttcagccac | ttctgcccac | ccaacgttgg | gccaaacaaa | aggagtccag | gccagtcaag | 4380 |
| atcgaaagat | acctccccca | attggaacag | agagactggc | cgaattcgg | caaggagggt | 4440 |
| ctgttgcaaca | agccccggcg | gggaccagtt | ttgtcgtctc | cgttggacac | agtggaaatc | 4500 |
| ggtcatttgg | tgtcaatgct | gtgtcagaag | gcttatcagg | ttggtcgcaa | tctgtgatgg | 4560 |
| ggaaccatcc | aatgcatcaa | caattatcag | acccaagcac | attctoccaa | catcagccaa | 4620 |
| tggagagaga | tgattctgga | atggtagccc | cctctaaca | ttttcatcag | cctatggcaa | 4680 |
| gtggttttgt | ggatttttct | aaaggtctgc | caatttccat | gtatggaggc | accataatac | 4740 |
| cctctcatcc | tcagcttgct | gatgttccag | gagggcctct | gtttaatgga | cttcacaatc | 4800 |
| cagatcctgc | ttggaacctt | atgataaaag | ttatocaaaa | ttcaactgaa | tgcaactgatg | 4860 |
| cccagcagat | ttggcctggc | acgtgggcac | ctcatattgg | aaacatgcac | ctcaaatatg | 4920 |
| tcaactaagt | tagaaggctc | ttactcttta | gccttggtta | agaaacctat | gaccttggaa | 4980 |
| gaaccatggg | gatttttttt | taatgtgcct | aagaaatttt | ctctgaggct | ttagcaatgg | 5040 |
| aaatttgatt | gcccattgta | taagaacaaa | ttgatttctc | atccacctga | ttatgttctc | 5100 |
| tggtttagttt | agccattttg | aacttaagat | catatgacct | tagtgctttt | ggctaacaac | 5160 |
| acagaatact | acttgtatgc | agaagagaat | tagttgatta | catgtttcaa | ccttttaggg | 5220 |
| tgataaatac | atgtataatt | gtttacatac | ttaaaaggaa | aaagttgagt | aaatttcttg | 5280 |
| tcatatagtg | gctctacgta | atgtagcctg | tattaatgtg | aaatatattac | cagaatatctc | 5340 |
| aataaaaaga | tgaacagtct | ttaaaaaaa | aaa | | | 5373 |

<210> 66

<211> 4333

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2772751CB1

<400> 66

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|------|
| gagaggggaca | gaggctggag | aaggatgtat | ggcctgccct | gggcttgtct | gttccctcct | 60 |
| gagcctgagc | cccttacctt | cctgacccca | tgaagcacac | actggctctg | ctggctcccc | 120 |
| tgctgggcct | gggcctgggg | ctggccctga | gtcagctggc | tgcagggggc | acagactgca | 180 |
| agttccttgg | cccgccagag | cacctgacat | tcacccagc | agccagggcc | cgggtggtgg | 240 |
| cccccgagt | tcgtgcgcca | ggactcctgg | actccctcta | tggcacctg | cgccgcttcc | 300 |
| tcctgggtgg | gcagctcaat | cctttccctt | cagagttgg | aaaggcccta | ctgaatgagc | 360 |
| tggcctccgt | gaaggagaat | gaggtggtgc | ggtacgaggc | gggctacgtg | gtatgcgtg | 420 |
| tgatcgccgg | cctctacctg | ctgctggtgc | ccactgccgg | gctttgcttc | tgctgtgccc | 480 |
| gctgccaccg | gcgtgcgggg | ggacgagtga | agacagagca | caaggcgctg | gcctgtgagc | 540 |
| gcgcggccct | catggtcttc | ctgctgtgta | ccaccctctt | gctgctgatt | ggtgtggtct | 600 |
| gtgcctttgt | caccaaccag | cgcacgcagt | aacagatggg | ccccagcatc | gaggccatgc | 660 |
| ctgagaccct | gctcagcctc | tggggcctgg | tctctgatgt | ccccaagag | ctgcaggccg | 720 |
| tggcacagca | attctccctg | ccccaggagc | aagtctcaga | ggagctggat | ggtgttggtg | 780 |
| tgagcattgg | gagcgcgcatc | cacactcagc | tcaggagctc | cgtgtacccc | ttgctggcgg | 840 |
| ccgtggggc | tttggggcag | gtcctgcagg | tcctcggtga | ccactgcaa | accttgaatg | 900 |
| ctacagtggt | agagctgcag | gcccggcagc | aggacctgga | gccagccatc | cgggaacacc | 960 |
| gggaccgcct | ccttgagctg | ctgcaggagg | ccagggtgcca | gggagattgt | gcaggggccc | 1020 |
| tgagctgggc | ccgcaccctg | gagctgggtg | ctgacttcag | ccagggtgcc | tctgtggacc | 1080 |
| atgtcctgca | ccagctaaaa | ggtgtccccg | aggccaactt | ctccagcatg | gtccaggagg | 1140 |
| agaacagcac | cctcaacgcc | cttccagccc | ttgctgccat | gcagacatcc | agcgtgtgtg | 1200 |
| aagagctgaa | gaaggcagtg | gcccagcagc | cggaaggggt | gaggacactg | gctgaagggt | 1260 |
| tcccgggctt | ggaggcagct | tcccgtctgg | cccaggcact | gcaggagggt | gaggagagca | 1320 |
| gcccggccca | cctgcaggag | gtgcagagat | acgagacctc | caggtggatc | gtgggctgcg | 1380 |
| tgctgtgtgc | cgtgggtccta | ttcgtgggtc | tctgcaacct | gctggggcctc | aatctgggca | 1440 |
| tctggggcct | gtctgccagg | gacgacccca | gccaccacga | agccaagggc | gaggctggag | 1500 |
| cccgttctct | catggcaggt | gtggggcctc | gcttccctct | tgctgcaccc | ctcatctctc | 1560 |
| tggtgttcgc | caccttctct | gtgggtggca | acgtgcagac | gctggtgtgc | cagagctggg | 1620 |
| agaacagcga | gctctttgag | tttgcaacca | ccccagggaa | cctgcccccg | tccatgaacc | 1680 |
| tgctcgcaat | tcttggcctg | aggaagaaca | tcagcatcca | ccaagcctat | cagcagtgc | 1740 |
| aggaaggggc | agcgtctctg | acagtcctgc | agctcaacga | ctcctacgac | ctggaggagc | 1800 |
| acctggatat | caaccagtat | accaacaagc | tacggcagga | gttgacagagc | ctgaaagtag | 1860 |
| acacacagag | cctggacctg | ctgagctcag | ccgcccgcgc | ggacctggag | gccctgcaga | 1920 |
| gcagtgggct | tcagcgcac | cactaccccg | acttccctct | tcagatccag | aggcccggtg | 1980 |
| tgaagaccag | catggagcag | ctggcccagg | agctgcaagg | actggcccag | gcccagaca | 2040 |
| attctgtgct | ggggcagcgg | ctgcaggagg | aggcccaagg | actcagaaac | cttcaccagg | 2100 |
| agaaggtcgt | ccccagcag | agccttctgt | caaagctcaa | cctcagcgtc | agggccctgg | 2160 |
| agtccctctg | cccgaatctc | cagctggaga | cctcagatgt | cctagccaat | gtcacctacc | 2220 |
| tgaagggaga | gctgcctgcc | tgggcagcca | ggatcctgag | gaatgtgagt | gagtgtttcc | 2280 |
| tggcccggga | gatgggttac | ttctcccagt | acgtggcctg | ggtgagagag | gagggtgactc | 2340 |
| agcgcattgc | cacctgccag | ccccctctcg | gagccctgga | caacagccgt | gtgatcctgt | 2400 |
| gtctcatgat | ggctgacccc | tgggaatgcc | tctggttctg | cctggcatgg | tgcaccttct | 2460 |
| tctgatcccc | cagcatcatc | tttgccgtca | agacctccaa | atacttccgt | cctatccgga | 2520 |
| aacgcctcag | ctccaccagc | tctgaggaga | ctcagctctt | ccacatcccc | cgggttacct | 2580 |
| ccctgaagct | gtagggcctt | gtggggtgag | gtgacctga | ggctgcctgt | cctccccctt | 2640 |
| gatttagcct | gggccacagg | acttcggtag | ctcttgcccc | agagcccagg | ctggcatcca | 2700 |
| ggcctgggact | gtcccaggtt | ccggttacc | tggccccacc | ttgcctgctc | ctttccaccc | 2760 |
| ctttctgctc | acgaccccc | tcattcacgc | tcagaatcac | atgggacttc | tgtgcagctg | 2820 |
| cagagccagc | aagtccctcc | aggtgtcacc | ccttaccccc | atgctgggtg | catcctcaca | 2880 |
| ggaagagcct | gttctccacc | tgctggagcc | tggaccctgg | ggtgggacag | aggcctcgte | 2940 |
| caaccccact | ccccctccg | tgtgtcttcc | ccctgccaa | cctccccctg | ccaagcctcc | 3000 |
| ccctgcccc | ctctgagccc | ctcgcccccc | acaccgtcct | catctggcct | cccccttggc | 3060 |
| ccccacttcc | ctcttatgcc | cttccctggc | ctttgcttcc | tccttagtgc | ccctcttcac | 3120 |
| catactccca | ctgctacctt | gctggcccca | gagaccaccc | tgcccaacca | aaccactcag | 3180 |
| gtaacgccac | taatcaggca | ggggccacca | tggcctaggt | ctgggctggc | tgcaggccct | 3240 |
| gcctcatggc | ctctgagccc | tccactgccc | cagggccttg | ggccctctgc | agatctcatc | 3300 |
| caggatttat | tgggtgtccag | tggggtgagg | gaggcctgtc | tgaaggccga | gcctccctgc | 3360 |
| ctgcacccaa | gttagaaaatg | gggggtaccag | cacttagctt | ctctctgagt | gctggctccc | 3420 |
| aaggaagggg | cctgggacct | ggggccacagt | gggggcttgc | ccttacctct | tcagaaggaa | 3480 |
| gcctcttcca | cagcccccac | ccaactttct | tcaggagtgt | ctgggtggca | gaacaggatt | 3540 |
| ttgcacggcc | ccttttatcc | tgcgcagtgt | gcctagggtc | atccccagcc | catccctgtg | 3600 |
| tcagccctga | gtgctggaca | ctgcgttcca | gaaatgagga | agaggagaga | gaagagatgg | 3660 |
| acagacctca | gatecattaa | agtgttctca | cttccctgag | acttggttct | gggtccctaa | 3720 |

| | | | | | | |
|------------|-------------|-------------|------------|------------|-------------|------|
| aaccaggttt | cctaggtctgg | gaccctgtac | atagttggtg | tttaatgagt | gtttatggag | 3780 |
| aggagagttc | taaggtcacc | tctggctgca | ggcatccagg | gattattcca | gcaatctgca | 3840 |
| ggtagggagt | gggtcccagc | ctgggagcct | gctgtcagga | gcaggcagac | ctggactcac | 3900 |
| agcctggctg | tgatgcttgt | tcgctcagct | tctccattta | tgagatgggg | agaatagtca | 3960 |
| cagcctcctc | aaagggttgt | gaaaatcaaa | tgtgataatt | tgtggaaagc | ccttagcagt | 4020 |
| ggcctggcac | aaaacaaatg | ctcagtggat | ggaagctgcc | tattattatt | gtcgttgttg | 4080 |
| ttgtttgcca | tgactgctct | gggccggggg | tagagctagc | atccgggcat | gtacgagggg | 4140 |
| agagggaggg | aggcctctat | tcaaaggcag | aaattccttt | aagatttgtg | tctgctgggt | 4200 |
| ttcagggagt | gtctgtgttg | tttgtttttg | tttgtttgtt | tgttttgaga | caggggtctcg | 4260 |
| ctctgtcacc | caggctggag | tgtagtgggtg | cagtcttggc | tcactgcaac | ctccacctcc | 4320 |
| tgggctcagc | gat | | | | | 4333 |

<210> 67

<211> 2213

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2793768CB1

<400> 67

| | | | | | | |
|-------------|-------------|------------|-------------|------------|------------|------|
| cggacctccc | gcgcgccccg | cacccgaccg | gctcagccgc | cggcagcgta | acacgcccta | 60 |
| cgctcgcttg | ctcgccggcc | tcagggcagg | caggcgggcg | cgggagaccc | cgccggggcc | 120 |
| gagacttggg | gcgggcgacg | aggaccaggt | tacggcctcc | tcgccatgtc | ctcggcctgc | 180 |
| gacgcggggc | accactaccc | cctgcacctc | ctagtctgga | aaaacgacta | ccggcagctc | 240 |
| gagaaggagg | tgcaaggcca | gaatgtggag | gctgtggacc | cacgaggtcg | aacattattg | 300 |
| catcttgctg | tttccttggg | acatttggaa | tctgctcgag | tcttactccg | acataaagca | 360 |
| gatgtgacaa | aagaaaaatc | ccagggtatg | acagttttac | atgaggctgt | gagcactggc | 420 |
| gatcctgaga | tgggtgtacac | agttctccaa | catcgagact | accacaacac | atccatggcc | 480 |
| cttgagggag | ttcctgagct | gtcctcaaaa | attctcgagg | ctccggattt | ctatgtgcag | 540 |
| atgaaatggg | aattcaccag | ctgggtgccc | ttggtttcta | gaatatgccc | aaatgatgtc | 600 |
| tgtcgcatct | ggaaaagtgg | tgccaaactg | cgcgctcgata | tcacattgct | gggatttgaa | 660 |
| aacatgagct | ggataagagg | gaggcgtagt | tttatattta | agggagaaga | caactgggcg | 720 |
| gagttaaatg | aagtcaacca | tgatgacaaa | gtggtcacca | ccgaacgctt | cgacctttcc | 780 |
| caagaaatgg | agcgctcac | tctggacttg | atgaagccaa | aaagcagggg | agttgagcgg | 840 |
| cggtccacaa | gccctgtcat | taacaccagc | ctcgatacta | aaaatatgtc | ttttgaaaga | 900 |
| actaaatccg | gattctgggg | ctggaggaca | gataaagcag | aagttgttaa | tggttacgaa | 960 |
| gcaaagggtt | acacagtaaa | caatgtgaat | gtgatcacca | aaatacgcac | agaacatctg | 1020 |
| accgaggagg | aaaaaaagag | atataaagca | gacaggaacc | cgctggaatc | tttgcgtggg | 1080 |
| actgtggaac | accaatttgg | tgacaaagg | gactccacca | cggaaatgtg | tactgcaaac | 1140 |
| aacccacag | ccatcacgcc | tgatgagtac | ttcaatgaag | agtttgatct | gaaagacagg | 1200 |
| gacattggaa | ggccgaaaga | gctgacgatt | agaacacaga | agtttaaagc | aatgttgttg | 1260 |
| atgtgtgaag | agtttccctt | ctctctgggt | gagcaggtca | ttcccatcat | tgacctaatg | 1320 |
| gctcgaacga | gtgctcattt | tgcaagactg | agagatttca | tcaaattgga | attcccacct | 1380 |
| ggatttccctg | tcaaaaataga | aattcccttg | tttcatgtct | taaatgcacg | gattacattt | 1440 |
| ggaaatgtta | atggctgtag | cactgccgaa | gaatctgtat | ctcaaatgtt | ggaagggacc | 1500 |
| caggctgatt | cagcttccca | catcacaaac | tttgagggtt | atcaatctgt | gtttgaaatt | 1560 |
| cccgaatctt | actatgttca | agacaatggc | agaaatgtgc | atttgcaaga | tgaagattac | 1620 |
| gagataatgc | agtttgccat | ccagcaaagt | ctgctggagt | ccagcaggag | ccaggaactt | 1680 |
| tcaggaccag | cttcgaatgg | agggatcagc | cagacaaaca | cctatgacgc | ccagtatgag | 1740 |
| agggccatcc | aggagagcct | cctcaccagc | acagaaggcc | tgtgccccag | cgccctgagc | 1800 |
| gagacaagcc | gttttgataa | tgacttgcag | ctagccatgg | agctctctgc | caaagagctg | 1860 |
| gaggaatggg | agctccggct | ccaggaggaa | gaggctgagc | tccagcaagt | cttacagctg | 1920 |
| tcactcactg | acaaatagac | ctttcagcct | gtgagcctct | gcacaaagca | gaggctgtgg | 1980 |
| gctgtcacag | atgctgtgtc | aaccagggcc | ctagggctaa | gggcctgcac | cttgctgtga | 2040 |
| tgacgcaggg | aacaactgcc | ccttctttat | gcagaggtgc | agaaccaggg | actcatgggg | 2100 |
| ccatccaggc | tgtcccttgg | cgtggagaag | ggaccaggga | ttgcaggccc | catctccagg | 2160 |
| ctaaaggggag | gagagcatca | tcactttcca | ttagctgtat | tggcttgagc | gtc | 2213 |

<210> 68

<211> 1142

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3035248CB1

<400> 68

| | | | | | | |
|-------------|------------|------------|------------|------------|-------------|------|
| gggctggcga | tccggatccc | gcaggcgcgc | tggetgcgct | gcccggtgt | ctgtcgatcat | 60 |
| ggtggggccc | tgggtgtatc | tggtggcggc | agttttgctc | atcggcctga | tcctcttcct | 120 |
| gactcgagc | cggggtcggg | cggcagcagc | tgacggagaa | ccactgcaca | atgaggaaga | 180 |
| gagggcagga | gcaggccagg | taggccgctc | tttgcccag | gagctgaag | aacagagaac | 240 |
| tgggaagcaga | ccccggcgtc | ggagggactt | gggcagccgt | ctacaggccc | agcgtcgagc | 300 |
| ccagcgagt | gcctgggaag | acggggatga | gaatgtgggt | caaactgtta | ttccagccca | 360 |
| ggaggaagaa | ggcattgaga | agccagcaga | agttcaccca | acagggaaaa | ttggagccaa | 420 |
| gaaactacgg | aagctagagg | aaaaacaggc | tcgaaaggct | cagcgagagg | cagaggaggc | 480 |
| tgaacgtgaa | gaacggaaac | gcctagagtc | ccaacgtgag | gccgaatgga | agaaggaaga | 540 |
| ggaacggctt | cgctgaagg | agaacagaa | ggaggaggaa | gagagggaag | ctcaggagga | 600 |
| gcaggcccgg | cgggatcacg | aggagtacct | gaaactgaag | gaggccttcg | tggtagaaga | 660 |
| agaagggtgtt | agcgaacca | tgactgagga | gcagtctcac | agcttcctga | cagaattcat | 720 |
| caattacatc | aagaagtcca | aggttgtgct | tttggaagat | ctggctttcc | agatgggcct | 780 |
| aaggactcag | gacgccataa | accgcatacc | ggacctgctg | acggagggga | ctctaacagg | 840 |
| tgtgattgac | gaccggggca | agtttatcta | cataaccca | gaggaactgg | ctgccgtggc | 900 |
| caatttcac | cgacagcggg | gccgggtgtc | catcacagag | cttgcccagg | ccagcaactc | 960 |
| cctcatctcc | tggggccagg | acctccctgc | ccaggcttca | gcctgactcc | agtccttcct | 1020 |
| tgagtgtatc | ctgtggccta | catgtgtctt | catccttccc | taatgccgtc | ttggggcagg | 1080 |
| gatggaatat | gaccagaaag | ttgtggattt | aaaggcctgt | gaatactgaa | aaaaaaaaaa | 1140 |
| gg | | | | | | 1142 |